

EXISTENCE AND UNIQUENESS RESULTS FOR BSDE WITH JUMPS: THE WHOLE NINE YARDS

ANTONIS PAPAPANTOLEON, DYLAN POSSAMAÏ, AND ALEXANDROS SAPLAOURAS

ABSTRACT. This paper is devoted to obtaining a wellposedness result for multidimensional BSDEs with possibly unbounded random time horizon and driven by a general martingale in a filtration only assumed to satisfy the usual hypotheses, *i.e.* the filtration may be stochastically discontinuous. We show that for stochastic Lipschitz generators and unbounded, possibly infinite, time horizon, these equations admit a unique solution in appropriately weighted spaces. Our result allows in particular to obtain a wellposedness result for BSDEs driven by discrete-time approximations of general martingales.

1. INTRODUCTION

A generally acknowledged fact is that backward stochastic differential equations (BSDEs for short) were introduced in their linear version by Bismut [16, 17] in 1973, as an adjoint equation in the Pontryagin stochastic maximum principle. However, around the same time, and most probably a bit before¹, Davis and Varaiya [45] (see in particular their Theorem 5.1) also studied what can be considered as a prototype of a linear BSDE for characterizing the value function and the optimal controls of stochastic control problems with drift control only. Such linear BSDEs, still in the context of the stochastic maximum principle, were also used by Arkin and Saksonov [1], Bensoussan [15] and Kabanov [74]. The first non-linear versions of these objects were once again introduced, under the form of a Riccati equation, by Bismut [18] and a few years later by Chitashvili [38] and Chitashvili and Mania [39, 40]. Nonetheless, the first study presenting a systematic treatment of non-linear BSDEs is the seminal paper by Pardoux and Peng [101]. Since then, and especially following the illuminating survey article of El Karoui, Peng and Quenez [53], BSDEs have become a particularly active field of research, due to their numerous potential applications to mathematical finance, partial differential equations, game theory, economics, and more generally in stochastic calculus and analysis².

Let $T > 0$ be fixed and consider a fixed filtered probability space $(\Omega, \mathcal{G}, \mathbb{G} := (\mathcal{G}_t)_{0 \leq t \leq T}, \mathbb{P})$ where \mathbb{G} is a Brownian filtration generated by some d -dimensional Brownian motion W . Solving a BSDE with terminal condition ξ (which is an \mathbb{R} -valued and \mathcal{G}_T -measurable random variable) and \mathbb{G} -adapted generator $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$, amounts to finding a pair of processes (Y, Z) which are respectively \mathbb{G} -progressively measurable and \mathbb{G} -predictable such that

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s^\top dW_s, \quad t \in [0, T],$$

holds, \mathbb{P} -a.s. After the work [101] obtained existence and uniqueness of the solution of the above BSDE in \mathbb{L}^2 -type spaces under square integrability assumptions on ξ and $f(s, 0, 0)$, and uniform Lipschitz continuity of f in (y, z) , generalizations of the theory have followed several different paths.

2010 *Mathematics Subject Classification.* 60G48, 60G55, 60G57, 60H05.

Key words and phrases. BSDEs, processes with jumps, stochastically discontinuous martingales, random time horizon, stochastic Lipschitz generator.

Alexandros Saplaouras gratefully acknowledges the financial support from the DFG Research Training Group 1845 “Stochastic Analysis with Applications in Biology, Finance and Physics”. Dylan Possamaï gratefully acknowledges the financial support of the ANR project PACMAN, ANR-16-CE05-0027. Moreover, all authors gratefully acknowledge the financial support from the PROCOPE project “Financial markets in transition: mathematical models and challenges”.

¹The authors are indebted to Saïd Hamadène for pointing out this reference. The published version of [45] states that the article was received on October 27, 1971. It is also present in bibliography of [17], though it is never referred to in the text.

²We emphasize that the references given below are just the tip of the iceberg, though most of them are, in our view, among the major ones of the field. Nonetheless, we do not make any claim about comprehensiveness of the following list.

The first one mainly aimed at weakening the Lipschitz assumption on f , and still considered Brownian filtrations. Hence, Mao [95] considered uniformly continuous generators, Hamadène [61] extended the result to the locally Lipschitz case, Lepeltier and San Martín [86] to the continuous and linear growth case in (y, z) , Briand and Carmona [22] to the case of a continuous generator, Lipschitz in z , with polynomial growth in y , Pardoux [100] to the case of a generator monotonic with arbitrary growth in y and Lipschitz in z . Some authors also obtained wellposedness results in \mathbb{L}^p -type spaces, among which we mention [53] for $p \geq 2$, Briand, Delyon, Hu, Pardoux and Stoica [25] and Briand and Hu [29] for $p \geq 1$ (see also the paper of Fan [57] for recent results and other references). Some attention has also been given to the so-called stochastic Lipschitz case, where the generator is Lipschitz continuous in (y, z) but with constants which are actually random processes themselves. There are few papers going in this direction, among which we can mention El Karoui and Huang [51], Bender and Kohlmann [14], Wang, Ran and Chen [111] as well as Briand and Confortola [23].

The first results going beyond the linear growth assumption in z , which assumed quadratic growth, were obtained by Kobylanski [79, 80, 81] for bounded ξ and f Lipschitz in y . Her results were then improved by Lepeltier and San Martín [87, 88], Briand, Lepeltier and San Martín [31] and revisited by Briand and Élie [27], but still for bounded ξ . Wellposedness in the quadratic case when ξ has sufficiently large exponential moments was then investigated by Briand and Hu [29, 30], followed by Delbaen, Hu and Richou [48, 49]. A specific quadratic setting with only square integrable terminal conditions has also been considered recently by Bahlali, Eddahbi and Ouknine [5]. A result with logarithmic growth was also obtained by Bahlali and El Asri [6]. The case of a generator with superquadratic growth in z was proved to be essentially ill-posed by Delbaen, Hu and Bao [47] in a general non-Markovian framework, before Richou [104], Cheridito and Stadje [37] and Masiero and Richou [96] proved that wellposedness could be recovered in a Markovian setting, when f has polynomial growth in (y, z) . Let us also mention the contributions by Cheridito and Nam [35], when ξ has a bounded Malliavin derivative, and by Drapeau, Heyne and Kupper [50] who considered minimal supersolutions of BSDEs when the generator is monotone in y and convex in z .

Most of the papers mentioned above treated the so-called one-dimensional BSDEs, that is for which the process Y is \mathbb{R} -valued, but extensions to multidimensional settings were also explored. Hence in Lipschitz or locally Lipschitz settings with monotonicity assumptions, we can mention the works of Pardoux [100], Bahlali [2, 3], Bahlali, Essaky, Hassani and Pardoux [9] and Bahlali, Essaky and Hassani [7, 8]. An early result in the case of a continuous generator in a Markovian setting was also treated by Hamadène, Lepeltier and Peng [62]. The quadratic multidimensional case is much more involved. Tevzadze [110] was the first to obtain a wellposedness result in the case of a bounded and sufficiently small terminal condition. It was then proved by Frei and dos Reis [59] and Frei [58] (see also Espinosa and Touzi [56] for a related problem) that even in seemingly benign situations, existence of global solutions could fail. Later on, Cheridito and Nam [36], Kardaras, Xing and Žitković [76], Kramkov and Pulido [82, 83], Hu and Tang [67], Jamneshan, Kupper and Luo [71], or more recently Kupper, Luo and Tangpi [85] and Élie and Possamaï [54] all obtained some results, but only in particular instances. Recently, Xing and Žitković [112] obtained quite general existence and uniqueness results, but in a Markovian framework, while Harter and Richou [63] and Jamneshan, Kupper and Luo [72] have obtained positive results in the general setting.

A second possible generalisation of these results consisted in extending them to the case where T is assumed to be a, possibly unbounded, stopping time. Hence, Peng [102], Darling and Pardoux [44], Briand and Hu [28], Royer [105], Hu and Tessitore [68] and Briand and Confortola [24] all studied this problem, applying it to homogenisation or representation problems for elliptic PDEs and stochastic control in infinite horizon.

Another avenue of generalisation concerned the underlying filtration itself, which could be assumed to no longer be Brownian, as well as the driving martingale, which could also be more general than a Brownian motion. In such cases, the predictable (martingale) representation property may fail to hold, and one has in general to add another martingale to the definition of a solution. Hence, for a given

martingale M , the problem becomes to find a triplet of processes (Y, Z, N) such that N is orthogonal to M and

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) dC_s - \int_t^T Z_s^\top dM_s - \int_t^T dN_s, \quad t \in [0, T], \quad \mathbb{P} - a.s.,$$

where the non-decreasing process C is absolutely continuous with respect to $\langle M \rangle$.

As far as we know, the first paper where such BSDEs appeared is due to Chitashvili [38] (see in particular the corollary at the end of page 91). Then, results on BSDEs driven by a general càdlàg martingale were obtained by Buckdahn [32], El Karoui *et al.* [53], as well as El Karoui and Huang [51], Briand, Delyon and Mémin [26] and Carbone, Ferrario and Santacrose [34], in Lipschitz type settings. The case of generators with quadratic growth has also been investigated by Tevzadze [110], Morlais [98], Réveillac [103], Imkeller, Réveillac and Richter [69], Mocha and Westray [97] and Barrieu and El Karoui [12]. More general versions of these equations, coined semi-martingale BSDEs, were also studied in depth in the context of financial applications, especially utility maximisation, see Bordigoni, Matoussi and Schweizer [20], as well as Hu and Schweizer [66], and mean-variance hedging, see Bobrovnytska and Schweizer [19], Mania and Tevzadze [92, 93, 94], Mania, Santacrose and Tevzadze [89, 90], Mania and Schweizer [91] as well as Jeanblanc, Mania, Santacrose and Schweizer [73].

When one has more information on the filtration, it may be possible to specify the orthogonal martingale N in the definition of the solution. For instance, if the filtration is generated by a Brownian motion and an orthogonal Poisson random measure, one ends up with the so-called BSDEs with jumps, which were introduced first by Tang and Li [109], followed notably by Buckdahn and Pardoux [33], Barles, Buckdahn and Pardoux [11], Situ [108], Royer [106], Becherer [13], El Karoui, Matoussi and Ngoupeyou [52], Kazi-Tani, Possamaï and Zhou [77, 78] and Fujii and Takahashi [60], while the specific case of Lévy processes was treated by Nualart and Schoutens [99] and later Bahlali, Eddahbi and Essaky [4]. A general presentation has been proposed recently by Kruse and Popier [84], to which we refer for more references.

One point that is actually shared by all the above references, is that the underlying filtration is assumed to be quasi-left continuous, which for instance rules out the possibility that any of the involved processes has jumps at predictable, and *a fortiori* deterministic times. The important simplification that arises is that the process C is then necessarily continuous in time. As far as we know, the first articles that went beyond this assumption were developed in a very nice series of papers by Cohen and Elliott [41] and Cohen, Elliott and Pearce [42], where the only assumption on the filtration is that the associated \mathbb{L}^2 space is separable, so that a very general martingale representation result due to Davis and Varaiya [46], involving countably many orthogonal martingales, holds. In these works, the martingales driving the BSDE are actually imposed by the filtration, and not chosen *a priori*, and the non-decreasing process C is not necessarily related to them, but has to be deterministic and can have jumps in general, though they have to be small for existence to hold (see [41, Theorem 5.1]). A similar approach is taken by Hassani and Ouknine in [64], where a form of BSDE is considered using generic maps from a space of semimartingales to the spaces of square-integrable martingales and of finite-variation processes integrable with respect to a given continuous increasing process. Similarly, Bandini [10] obtained wellposedness results in a context of a general filtration allowing for jumps, with a fixed driving martingale and associated random process C , which must have again small jumps, see [10, Equation (1.1)]. Let us also mention the work by Confortola, Fuhrman and Jacod [43] which concentrates on the pure-jump general case and gives in particular counterexamples to existence. Finally, Bouchard, Possamaï, Tan and Zhou [21] provide a general method to obtain *a priori* estimates in general filtrations when the martingale driving the equation has quadratic variation absolutely continuous with respect to the Lebesgue measure.

In this paper, we improve the general result on existence and uniqueness of solutions of backward stochastic differential equations given by El Karoui and Huang in [51] to the case where the martingale M driving the equation is possibly stochastically discontinuous. In other words, our framework includes as driving martingales discrete-time approximations of general martingales as well as

K —almost quasi-left-continuous martingales, *i.e.* processes whose compensator has jumps which are almost surely bounded by some constant K . Unlike all the related papers mentioned above (with the notable exception of [41], see their Theorem 6.1, albeit with a deterministic Lipschitz constant), this bound K can actually be arbitrarily large. However, the product of this bound and the maximum (of functionals) of the Lipschitz constants needs to be small, which is in line with the previous literature. Otherwise, we remain in the same relaxed framework regarding the generator, that is to say we assume that it satisfies a stochastic Lipschitz property, and do not assume that the martingale possesses the predictable representation property. Furthermore, we work in a setting with random horizon. This result enables us to treat under the same framework continuous-time as well as discrete-time BSDEs. The method of proof is somehow similar to the one given in [51], but the required estimates are much harder to prove in our setting due to the possible jumps of the non-decreasing process C . We also emphasise that this wellposedness result will be of fundamental importance in a related, forthcoming work, where we will use it to study robustness properties of general BSDEs, extending well-known results on stability of semimartingale decompositions with respect to the extended convergence.

This paper is structured as follows: in Section 2 we introduce the notation and several results that will be useful in the analysis. In Section 3 we prove *a priori* estimates for the considered class of BSDEs and provide the existence and uniqueness results. Finally, Section 4 discusses some applications of the main results, while the Appendices contains proofs and auxiliary results.

Notation. Let \mathbb{R}_+ denote the set of non-negative real numbers. For any positive integer ℓ , and for any $(x, y) \in \mathbb{R}^\ell \times \mathbb{R}^\ell$, $|x|$ will denote the Euclidean norm of x . For any additional integer q , a $q \times \ell$ —matrix with real entries will be considered as an element of $\mathbb{R}^{q \times \ell}$. For any $z \in \mathbb{R}^{q \times \ell}$, its transpose will be denoted by $z^\top \in \mathbb{R}^{\ell \times q}$. We endow $\mathbb{R}^{q \times \ell}$ with the norm defined for any $z \in \mathbb{R}^{q \times \ell}$ by $\|z\|^2 := \text{Tr}[z^\top z]$ and remind the reader that this norm derives from the inner product defined for any $(z, u) \in \mathbb{R}^{q \times \ell} \times \mathbb{R}^{q \times \ell}$ by $\text{Tr}[zu^\top]$. We abuse notation and denote by 0 the neutral element in the group $(\mathbb{R}^{q \times \ell}, +)$. Furthermore, for any finite dimensional topological space E , $\mathcal{B}(E)$ will denote the associated Borel σ —algebra. In addition, for any other finite dimensional space F , and for any non-negative measure ν on $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(E))$, we will denote indifferently Lebesgue-Stieltjes integrals of any measurable map $f : (\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(E)) \rightarrow (F, \mathcal{B}(F))$, by

$$\int_{(u,t] \times A} f(s, x) \nu(ds, dx) \text{ or } \int_u^t \int_A f(s, x) \nu(ds, dx), \text{ for any } (t, A) \in \mathbb{R}_+ \times \mathcal{B}(E),$$

$$\int_{(u, \infty) \times A} f(s, x) \nu(ds, dx) \text{ or } \int_u^\infty \int_A f(s, x) \nu(ds, dx), \text{ for any } A \in \mathcal{B}(E),$$

where the integrals are to be understood in a component-wise sense. Finally, we fix positive integers d, m and n throughout the rest of the paper.

2. PRELIMINARIES

2.1. The stochastic basis. Let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be a complete stochastic basis in the sense of [70, Definition I.1.3]. Expectations under \mathbb{P} will be denoted by $\mathbb{E}[\cdot]$. Let τ be a \mathbb{G} —stopping time; we will then denote the set of \mathbb{R}^m —valued and square-integrable \mathbb{G} —martingales stopped at τ ³ by \mathcal{H}^2 . Let $X \in \mathcal{H}^2$, then its norm will be defined by

$$\|X\|_{\mathcal{H}^2}^2 := \mathbb{E}[\text{Tr}[\langle X \rangle_\tau]].$$

In the sequel we will say that $M, N \in \mathcal{H}^2$ are (*mutually*) *orthogonal*, denoted by $M \perp N$, if MN is a martingale, see [70, Definition I.4.11.a, Lemma 4.13.c, Proposition I.4.15], and we define the space of martingales orthogonal to M by

$$\mathcal{H}^2(M^\perp) := \{N \in \mathcal{H}^2, \langle M, N \rangle = 0\}.$$

³As usual, for a martingale X , the corresponding martingale stopped at τ , denoted by X^τ , is defined by $X_t^\tau := X_{t \wedge \tau}$, $t \geq 0$.

A martingale $M \in \mathcal{H}^2$ will be called a *purely discontinuous* martingale if $M_0 = 0$ and if it is orthogonal to all continuous martingales of \mathcal{H}^2 . Using [70, Corollary I.4.16] we can decompose \mathcal{H}^2 as follows

$$\mathcal{H}^2 =: \mathcal{H}^{2,c} \oplus \mathcal{H}^{2,d},$$

where $\mathcal{H}^{2,c}$ is the subspace of \mathcal{H}^2 consisting of continuous square-integrable martingales and $\mathcal{H}^{2,d}$ is the subspace of \mathcal{H}^2 consisting of all purely discontinuous square-integrable martingales. It follows then from [70, Theorem I.4.18], that any \mathbb{G} -martingale $X \in \mathcal{H}^2$ admits a unique (up to \mathbb{P} -indistinguishability) decomposition

$$X = X_0 + X^c + X^d,$$

where $X_0^c = X_0^d = 0$. $X^c \in \mathcal{H}^{2,c}$ will be called the *continuous martingale part of X* and $X^d \in \mathcal{H}^{2,d}$ the *purely discontinuous martingale part of X* .

2.2. Stochastic integrals. Consider an arbitrary \mathbb{G} -stopping time τ and an arbitrary martingale $X \in \mathcal{H}^2$. We start by defining

$$\mathbb{H}^2(X^c) := \left\{ Z : (\Omega \times \mathbb{R}_+, \mathcal{P}) \longrightarrow (\mathbb{R}^{d \times m}, \mathcal{B}(\mathbb{R}^{d \times m})), \mathbb{E} \left[\text{Tr} \left[\int_0^\tau Z_t^\top Z_t d\langle X^c \rangle_t \right] \right] < \infty \right\},$$

where \mathcal{P} denotes the predictable σ -field of \mathbb{G} -adapted processes on $\Omega \times \mathbb{R}_+$.

Let $Z \in \mathbb{H}^2(X^c)$, then the Itô stochastic integral of Z with respect to X^c is well defined and is an element of $\mathcal{H}^{2,c}$, see [70, Theorem I.4.40, Property I.4.36]. It will be denoted by $\int_0^\cdot Z_s dX_s^c$ or $Z \cdot X^c$ interchangeably, and we will also use the same notation for any Stieltjes-type integral. Moreover, we have that $(Z^\top Z) \cdot \langle X^c \rangle = \langle Z \cdot X^c \rangle$, hence the following equality holds

$$\|Z\|_{\mathbb{H}^2(X^c)}^2 := \mathbb{E} \left[\text{Tr} \left[\int_0^\tau Z_t^\top Z_t d\langle X^c \rangle_t \right] \right] = \mathbb{E} [\text{Tr}[\langle Z \cdot X^c \rangle_\tau]].$$

We will denote the space of Itô stochastic integrals of processes in $\mathbb{H}^2(X^c)$ with respect to X^c by $\mathcal{L}^2(X^c)$, and we remind the reader that, by [70, Theorem I.4.40.b], $\mathcal{L}^2(X^c) \subset \mathcal{H}^2$.

Let us define the space

$$(\tilde{\Omega}, \tilde{\mathcal{P}}) := (\Omega \times \mathbb{R}_+ \times \mathbb{R}^n, \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n)).$$

A measurable function $U : (\tilde{\Omega}, \tilde{\mathcal{P}}) \longrightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is called $\tilde{\mathcal{P}}$ -*measurable function* or simply \mathbb{G} -*predictable function*.

Let $\mu := \{\mu(\omega; dt, dx)\}_{\omega \in \Omega}$ be a random measure on $\mathbb{R}_+ \times \mathbb{R}^n$, i.e. a family of non-negative measures defined on $(\mathbb{R}_+ \times \mathbb{R}^n, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^n))$ satisfying $\mu(\omega; \{0\} \times \mathbb{R}^n) = 0$, identically. For a \mathbb{G} -predictable function U , we define the process

$$U \star \mu(\omega) := \begin{cases} \int_{(0, \cdot] \times \mathbb{R}^n} U(\omega, s, x) \mu(\omega; ds, dx), & \text{if } \int_{(0, \cdot] \times \mathbb{R}^n} |U(\omega, s, x)| \mu(\omega; ds, dx) < \infty, \\ \infty, & \text{otherwise.} \end{cases}$$

We can associate to the purely discontinuous part X^d of X the \mathbb{G} -optional integer-valued random measure μ^X on $\mathbb{R}_+ \times \mathbb{R}^n$ defined by

$$\mu^X(\omega; dt, dx) := \sum_{s > 0} \mathbb{1}_{\{\Delta X_s(\omega) \neq 0\}} \delta_{(s, \Delta X_s(\omega))}(dt, dx),$$

see [70, Proposition II.1.16], where, for any $z \in \mathbb{R}_+ \times \mathbb{R}^n$, δ_z denotes the Dirac measure at the point z . Notice that $\mu^X(\omega; \mathbb{R}_+ \times \{0\}) = 0$. Moreover, for a \mathbb{G} -predictable stopping time σ we define the random variable

$$\begin{aligned} \int_{\mathbb{R}^n} U(\omega, \sigma, x) \mu^X(\omega; \{\sigma\} \times dx) &:= U(\omega, \sigma(\omega), \Delta X_{\sigma(\omega)}(\omega)) \mathbb{1}_{\{\Delta X_{\sigma} \neq 0, |U(\omega, \sigma(\omega), \Delta X_{\sigma(\omega)}(\omega))| < \infty\}} \\ &\quad + \infty \mathbb{1}_{\{|U(\omega, \sigma(\omega), \Delta X_{\sigma(\omega)}(\omega))| = \infty\}}. \end{aligned}$$

Since $X \in \mathcal{H}^2$, the compensator of μ^X under \mathbb{P} exists, see [70, Theorem II.1.8]. This is the unique, up to a \mathbb{P} -null set, \mathbb{G} -predictable random measure ν^X on $\mathbb{R}_+ \times \mathbb{R}^n$, for which

$$\mathbb{E}[U \star \mu_\infty^X] = \mathbb{E}[U \star \nu_\infty^X]$$

holds for every non-negative \mathbb{G} -predictable function U .

For a non-negative \mathbb{G} -predictable function U and a \mathbb{G} -predictable time σ , whose graph is denoted by $[\![\sigma]\!]$ (see [70, Notation I.1.22] and the comments afterwards), we define the random variable

$$\int_{\mathbb{R}^n} U(\omega, \sigma, x) \nu^X(\omega; \{\sigma\} \times dx) := \int_{\mathbb{R}_+ \times \mathbb{R}^n} U(\omega, \sigma(\omega), x) \mathbb{1}_{[\![\sigma]\!]} \nu^X(\omega; ds, dx),$$

if $\int_{\mathbb{R}_+ \times \mathbb{R}^n} |U(\omega, \sigma(\omega), x)| \mathbb{1}_{[\![\sigma]\!]} \nu^X(\omega; ds, dx) < \infty$, otherwise we define it to be equal to ∞ . By [70, Property II.1.11], we have

$$\int_{\mathbb{R}^n} U(\omega, \sigma, x) \nu^X(\omega; \{\sigma\} \times dx) = \mathbb{E} \left[\int_{\mathbb{R}^n} U(\omega, \sigma, x) \mu^X(\omega; \{\sigma\} \times dx) \middle| \mathcal{G}_{\sigma-} \right].$$

In order to simplify notation further, let us denote for any \mathbb{G} -stopping time σ

$$\widehat{U}_\sigma^X := \int_{\mathbb{R}^n} U(\omega, \sigma, x) \nu^X(\omega; \{\sigma\} \times dx).$$

In order to define the stochastic integral of a \mathbb{G} -predictable function U with respect to the *compensated integer-valued measure* $\tilde{\mu}^X := \mu^X - \nu^X$, we will need to consider the following class

$$G_2(\tilde{\mu}^X) = \left\{ U : (\tilde{\Omega}, \tilde{\mathcal{P}}) \longrightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)), \mathbb{E} \left[\sum_{t>0} \left(U(t, \Delta X_t) \mathbb{1}_{\{\Delta X_t \neq 0\}} - \widehat{U}_t^X \right)^2 \right] < \infty \right\}.$$

Any element of $G_2(\tilde{\mu}^X)$ can be associated to an element of $\mathcal{H}^{2,d}$ via

$$G_2(\tilde{\mu}^X) \ni U \longmapsto U \star \tilde{\mu}^X \in \mathcal{H}^{2,d},$$

see [70, Definition II.1.27, Proposition II.1.33.a] and [65, Theorem XI.11.21]. This is unique up to \mathbb{P} -indistinguishability, and we call $U \star \tilde{\mu}^X$ the *stochastic integral of U with respect to $\tilde{\mu}^X$* . We will also make use of the following notation for the space of stochastic integrals with respect to $\tilde{\mu}^X$ which are square integrable martingales

$$\mathcal{K}(\tilde{\mu}^X) := \left\{ U \star \tilde{\mu}^X, U \in G_2(\tilde{\mu}^X), U \star \tilde{\mu}^X \in \mathcal{H}^{2,d} \right\}.$$

Moreover, by [70, Theorem II.1.33] or [65, Theorem 11.21], we have

$$\mathbb{E}[\langle U \star \tilde{\mu}^X \rangle_\infty] < \infty \quad \text{if and only if} \quad U \in G_2(\tilde{\mu}^X),$$

which enables us to define the following more convenient space

$$\mathbb{H}^2(X^d) := \left\{ U : (\tilde{\Omega}, \tilde{\mathcal{P}}) \longrightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)), \mathbb{E} \left[\int_0^\tau d\text{Tr} [\langle U \star \tilde{\mu}^X \rangle_t] \right] < \infty \right\},$$

and we emphasise that we have the direct identification

$$\mathbb{H}^2(X^d) = G_2(\tilde{\mu}^X).$$

2.3. Suitable spaces and associated results. In the sequel we will make use of the following assumption, which is based on [70, Property II.1.2 and Proposition II.2.9]. Let us first define the maps $\mathbb{R}^n \ni x \xrightarrow{q} xx^\top \in \mathbb{R}^{n \times n}$ and $\mathbb{R}^n \ni x \xrightarrow{I} x \in \mathbb{R}^n$.

Assumption (C). A pair (X, C^X) of a semimartingale X and a predictable, càdlàg and increasing process C^X satisfies Assumption (C) if each component of $\langle X^c \rangle$ is absolutely continuous w.r.t. to C^X and if the disintegration property given C^X holds for the compensator ν^X . In other words, if there exists a predictable, positive definite and symmetric $m \times m$ -matrix $\frac{d\langle X^c \rangle}{dC^X}$ and a transition kernel

$K^X : (\Omega \times \mathbb{R}_+, \mathcal{P}) \longrightarrow \mathcal{R}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, where $\mathcal{R}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is the space of Radon measures on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, such that for any $1 \leq i, j \leq m$

$$\langle X^c \rangle^{ij} = \int_0^\cdot \frac{d\langle X^c \rangle_s^{ij}}{dC_s^X} dC_s^X \quad \text{and} \quad \nu^X(\omega; dt, dx) = K_t^X(\omega; dx) dC_t^X.$$

Remark 2.1. Let X be a semimartingale, then there exist several possible choices for C^X such that Assumption (C) is satisfied. In [70, Proposition II.2.9], for example, the following process is used

$$\tilde{C}^X := \sum_{i,j=1}^n \text{Var}(\langle X^{c,i}, X^{c,j} \rangle) + q \star \nu^X,$$

while one could also take

$$\overline{C}^X := \text{Tr}[\langle X^c \rangle] + q \star \nu^X.$$

Remark 2.2. Consider (X, C^X) that satisfies Assumption (C), then the Radon-Nikodým derivative $\frac{d\langle X^c \rangle}{dC^X}$ is \mathbb{G} -predictable, positive definite and symmetric. Indeed, the predictability follows from [70, Proposition I.3.13], while the symmetry is immediately inherited from the symmetry of $\langle X^c \rangle$. Moreover, the positive definiteness can be proved by following similar arguments to [70, Proposition II.2.9]. The above properties enable us to define the following \mathbb{G} -predictable process c^X

$$c^X := \left(\frac{d\langle X^c \rangle}{dC^X} \right)^{\frac{1}{2}}. \quad (2.1)$$

In addition, if we define the random measure μ_Δ^X on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, for any $t \geq 0$, via

$$\mu_\Delta^X(\omega; [0, t]) := \sum_{0 < s \leq t} (\Delta C_s^X(\omega))^2 \quad \text{then it holds} \quad \frac{d\mu_\Delta^X}{dC^X}(t) = \Delta C_t^X. \quad (2.2)$$

Assume now that X is a square-integrable martingale and there exists a process C^X such that (X, C^X) satisfies Assumption (C). The process X is not assumed quasi-left-continuous, hence it is possible that it has fixed times of discontinuities.

Using [70, Proposition II.2.29.b] we have that $X \in \mathcal{H}^2$ if and only if the following holds

$$\mathbb{E} \left[\langle X^c \rangle_\tau + \|q\|^2 \star \nu_\tau^X \right] < \infty.$$

The predictable quadratic variation of X admits then the following representation

$$\langle X \rangle_\cdot = \langle X^c \rangle_\cdot + q \star \nu^X - \sum_{0 < s \leq \cdot} \int_{\mathbb{R}^n} x \nu^X(\{s\} \times dx) \int_{\mathbb{R}^n} x^\top \nu^X(\{s\} \times dx).$$

Let U be a \mathbb{G} -predictable function taking values in $\mathbb{R}^{q \times \ell}$ for some positive integers q and ℓ . Then we define, abusing notation slightly,

$$\hat{K}_t^X(U_t(\omega; \cdot))(\omega) := \int_{\mathbb{R}^n} U_t(\omega; x) K_t^X(\omega; dx), \quad t \geq 0,$$

where K^X is the transition kernel from Assumption (C). Using Assumption (C) and (2.2), we get that

$$\int_{\mathbb{R}^n} U_t(\omega; x) \nu^X(\omega; \{t\} \times dx) = \hat{K}_t^X(U_t(\omega; \cdot))(\omega) \Delta C_t^X(\omega), \quad t \geq 0,$$

and

$$\begin{aligned} \int_{\mathbb{R}^n} |U_t(\omega; x)|^2 \nu^X(\omega; \{t\} \times dx) &= \int_{\mathbb{R}^n} |U_t(\omega; x)|^2 K_t^X(\omega; dx) \Delta C_t^X(\omega) \\ &= \hat{K}_t^X(|U_t(\omega; \cdot)|^2)(\omega) \Delta C_t^X(\omega), \quad t \geq 0. \end{aligned} \quad (2.3)$$

Using the previous definitions and results, we can rewrite $\langle X \rangle$ as follows

$$\begin{aligned}
\langle X \rangle &= \int_0^\cdot \left((c_s^X)^\top c_s^X + \int_{\mathbb{R}^n} x x^\top K_s^X(dx) \right) dC_s^X \\
&\quad - \sum_{0 < s \leq \cdot} \int_{\mathbb{R}^n} x K_s^X(dx) \Delta C_s^X \int_{\mathbb{R}^n} x^\top K_s^X(dx) \Delta C_s^X \\
&= \int_0^\cdot \left((c_s^X)^\top c_s^X + \widehat{K}_s^X(q) \right) dC_s^X - \sum_{s \leq \cdot} \widehat{K}_s^X(I) \left(\widehat{K}_s^X(I) \right)^\top (\Delta C_s^X)^2 \\
&= \int_0^\cdot \left((c_s^X)^\top c_s^X + \widehat{K}_s^X(q) - \Delta C_s^X \widehat{K}_s^X(I) \left(\widehat{K}_s^X(I) \right)^\top \right) dC_s^X.
\end{aligned}$$

We proceed now by defining the spaces that will be necessary for our analysis, see also [51]. Let $\beta \geq 0$ and $A : (\Omega \times \mathbb{R}_+, \mathcal{G} \otimes \mathcal{B}(\mathbb{R}_+)) \longrightarrow \mathbb{R}_+$ be a càdlàg, increasing and measurable process. We then define the following spaces, where the dependence on A is suppressed for ease of notation:

$$\begin{aligned}
\mathbb{L}_\beta^2(\mathcal{G}_\tau; \mathbb{R}^d) &:= \left\{ \xi, \mathbb{R}^d\text{-valued, } \mathcal{G}_\tau\text{-measurable, } \|\xi\|_{\mathbb{L}_\beta^2(\mathcal{G}_\tau; \mathbb{R}^d)}^2 := \mathbb{E} \left[e^{\beta A_\tau} |\xi|^2 \right] < \infty \right\}, \\
\mathcal{H}_\beta^2 &:= \left\{ M \in \mathcal{H}^2, \|M\|_{\mathcal{H}_\beta^2}^2 := \mathbb{E} \left[\int_0^\tau e^{\beta A_t} d\text{Tr}[\langle M \rangle_t] \right] < \infty \right\}, \\
\mathbb{H}_\beta^2(X) &:= \left\{ \phi \text{ is an } \mathbb{R}^d\text{-valued } \mathbb{G}\text{-optional semimartingale with càdlàg paths and} \right. \\
&\quad \left. \|\phi\|_{\mathbb{H}_\beta^2(X)}^2 := \mathbb{E} \left[\int_0^\tau e^{\beta A_t} |\phi_t|^2 dC_t^X \right] < \infty \right\}, \\
\mathcal{S}_\beta^2(X) &:= \left\{ \phi \text{ is an } \mathbb{R}^d\text{-valued } \mathbb{G}\text{-optional semimartingale with càdlàg paths and} \right. \\
&\quad \left. \|\phi\|_{\mathcal{S}_\beta^2(X)}^2 := \mathbb{E} \left[\sup_{t \in [0, \tau]} e^{\beta A_t} |\phi_t|^2 \right] < \infty \right\}, \\
\mathbb{H}_\beta^2(X^c) &:= \left\{ Z \in \mathbb{H}^2(X^c), \|Z\|_{\mathbb{H}_\beta^2(X^c)}^2 := \mathbb{E} \left[\int_0^\tau e^{\beta A_t} d\text{Tr}[\langle Z \cdot X^c \rangle_t] \right] < \infty \right\}, \\
\mathbb{H}_\beta^2(X^d) &:= \left\{ U \in \mathbb{H}^2(X^d), \|U\|_{\mathbb{H}_\beta^2(X^d)} < \infty, \right. \\
&\quad \left. \text{with } \|U\|_{\mathbb{H}_\beta^2(X^d)} := \mathbb{E} \left[\int_0^\tau e^{\beta A_t} d\text{Tr}[\langle U \star \tilde{\mu}^X \rangle_t] \right] \right\}, \\
\mathcal{H}_\beta^2(X^\perp) &:= \left\{ M \in \mathcal{H}^2(X^\perp), \|M\|_{\mathcal{H}_\beta^2(X^\perp)}^2 := \mathbb{E} \left[\int_0^\tau e^{\beta A_t} d\text{Tr}[\langle M \rangle_t] \right] < \infty \right\}.
\end{aligned}$$

Finally, for $(Y, Z, U, N) \in \mathbb{H}_\beta^2(X) \times \mathbb{H}_\beta^2(X^c) \times \mathbb{H}_\beta^2(X^d) \times \mathbb{H}_\beta^2(X^\perp)$ and assuming that $A = \alpha^2 \cdot C^X$ for a measurable process $\alpha : (\Omega \times \mathbb{R}_+, \mathcal{G} \otimes \mathcal{B}(\mathbb{R}_+)) \longrightarrow \mathbb{R}$, we define

$$\|(Y, Z, U, N)\|_{\beta, X}^2 := \|\alpha Y\|_{\mathbb{H}_\beta^2(X)}^2 + \|Z\|_{\mathbb{H}_\beta^2(X^c)}^2 + \|U\|_{\mathbb{H}_\beta^2(X^d)}^2 + \|N\|_{\mathbb{H}_\beta^2(X^\perp)}^2,$$

and for $(Y, Z, U, N) \in \mathcal{S}_\beta^2(X) \times \mathbb{H}_\beta^2(X^c) \times \mathbb{H}_\beta^2(X^d) \times \mathbb{H}_\beta^2(X^\perp)$, we define

$$\|(Y, Z, U, N)\|_{\star, \beta, X}^2 := \|Y\|_{\mathcal{S}_\beta^2(X)}^2 + \|Z\|_{\mathbb{H}_\beta^2(X^c)}^2 + \|U\|_{\mathbb{H}_\beta^2(X^d)}^2 + \|N\|_{\mathbb{H}_\beta^2(X^\perp)}^2.$$

The next lemma will be useful for future computations and, in addition, justifies the definition of the norms on the spaces provided above.

Lemma 2.3. *Let $(Z, U) \in \mathbb{H}_\beta^2(X^c) \times \mathbb{H}_\beta^2(X^d)$. Then*

$$\|Z\|_{\mathbb{H}_\beta^2(X^c)}^2 = \mathbb{E} \left[\int_0^\tau e^{\beta A_t} \|c_t Z_t\|^2 dC_t^X \right], \quad (2.4)$$

$$\|U\|_{\mathbb{H}_\beta^2(X^d)}^2 = \mathbb{E} \left[\int_0^\tau e^{\beta A_t} (\|U_t(\cdot)\|_t^X)^2 dC_t^X \right], \quad (2.5)$$

where for every $(t, \omega) \in \mathbb{R}_+ \times \Omega$

$$\left(\|U_t(\omega; \cdot)\|_t^X(\omega) \right)^2 := \widehat{K}_t^X(|U_t(\omega; \cdot)|^2)(\omega) - \Delta C_t^X(\omega) |\widehat{K}_t^X(U_t(\omega; \cdot))(\omega)|^2 \geq 0.$$

Furthermore

$$\|Z \cdot X^c + U \star \tilde{\mu}^X\|_{\mathcal{H}_\beta^2}^2 = \|Z\|_{\mathbb{H}_\beta^2(X^c)}^2 + \|U\|_{\mathbb{H}_\beta^2(X^d)}^2.$$

Proof. Let $Z \in \mathbb{H}_\beta^2(X^c)$, then using [70, Theorem III.6.4], we get that $Z \cdot X^c \in \mathcal{H}^{2,c}$ with

$$\langle Z \cdot X^c \rangle = Z \frac{d\mu_\Delta^X}{dC^X} Z^\top \cdot C^X = Z c^X (c^X)^\top Z^\top \cdot C^X, \quad (2.6)$$

for c^X as introduced in (2.1). The first result is then obvious.

Now, let $U \in \mathbb{H}_\beta^2(X^d)$, then similarly to the previous computations, we have

$$\langle U \star \tilde{\mu}^X \rangle = \int_0^\cdot \left(\widehat{K}_t^X((U_t U_t^\top)(\cdot)) - \widehat{K}_t^X(U_t(\cdot)) \left(\widehat{K}_t^X(U_t(\cdot)) \right)^\top \Delta C_s^X \right) dC_s^X, \quad (2.7)$$

from which the second result is also clear. Moreover, we have

$$\begin{aligned} \|Z \cdot X^c + U \star \tilde{\mu}^X\|_{\mathcal{H}_\beta^2}^2 &= \mathbb{E} \left[\int_0^\tau e^{\beta A_t} d\text{Tr} [\langle Z \cdot X^c + U \star \tilde{\mu}^X \rangle_t] \right] \\ &= \mathbb{E} \left[\int_0^\tau e^{\beta A_t} d\text{Tr} [\langle Z \cdot X^c \rangle_t] + \int_0^\tau e^{\beta A_t} d\text{Tr} [\langle U \star \tilde{\mu}^X \rangle_t] \right], \end{aligned}$$

where the second equality holds due to the orthogonality of X^c and X^d , which implies that we have $\langle Z \cdot X^c, U \star \tilde{\mu}^X \rangle = 0$, see [70, Proposition 4.15].

Notice finally that the process $\text{Tr}[\langle U \star \tilde{\mu}^X \rangle]$ is non-decreasing, and observe that

$$\Delta \text{Tr} [\langle U \star \tilde{\mu}^X \rangle_t] = \left(\|U_t(\cdot)\|_t^X \right)^2 \Delta C_t^X, \quad t \geq 0. \quad (2.8)$$

Since C^X is non-decreasing, we can deduce that $\|U_t(\cdot)\|_t^X \geq 0$. \square

We conclude this section with the following convenient result. Define the following space for $dC_t^X \otimes d\mathbb{P} - a.e. (t, \omega) \in \mathbb{R}_+ \times \Omega$

$$\mathfrak{H}_{t,\omega}^X := \overline{\left\{ \mathcal{U} : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \longrightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)), \|\mathcal{U}(\cdot)\|_t^X(\omega) < \infty \right\}}.$$

Define also

$$\mathfrak{H}^X := \left\{ U : [0, T] \times \Omega \times \mathbb{R}^n \longrightarrow \mathbb{R}^d, U_t(\omega; \cdot) \in \mathfrak{H}_{t,\omega}^X, \text{ for } dC_t^X \otimes d\mathbb{P} - a.e. (t, \omega) \in \mathbb{R}_+ \times \Omega \right\}.$$

Lemma 2.4. *The space $(\mathfrak{H}_{t,\omega}^X, \|\cdot\|_t^X(\omega))$ is Polish, for $dC_t^X \otimes d\mathbb{P} - a.e. (t, \omega) \in \mathbb{R}_+ \times \Omega$.*

Proof. The fact that $\|\cdot\|_t^X(\omega)$ is indeed a norm is immediate from (2.8) and Kunita–Watanabe’s inequality. Then, the space is clearly Polish since $K_t(\omega; dx)$ is a regular measure, as it integrates $x \mapsto |x|^2$ for $dC_t^X \otimes d\mathbb{P} - a.e. (t, \omega) \in \mathbb{R}_+ \times \Omega$. \square

2.4. A useful lemma for generalised inverses. In the following sections we will need a result on generalised inverses which is stated as a corollary of the following lemma. The proof is presented in Appendix A.

Lemma 2.5. *Let g be a non-decreasing sub-multiplicative function on \mathbb{R}_+ , that is to say*

$$g(x+y) \leq \ell g(x)g(y),$$

for some $\ell > 0$ and for every $x, y \in \mathbb{R}_+$. Let A be a càdlàg and non-decreasing function and define its left-continuous inverse L by

$$L_s := \inf \{t \geq 0, A_t \geq s\}.$$

Then it holds that

$$\int_0^t g(A_s) dA_s \leq \ell g \left(\max_{\{s, L_s < \infty\}} \Delta A_{L_s} \right) \int_{A_0}^{A_t} g(s) ds.$$

Corollary 2.6. *Let A and g as in Lemma 2.5 with the additional assumption that A has uniformly bounded jumps, say by K . Then there exists a universal constant $K' > 0$ such that*

$$\int_0^t g(A_s) dA_s \leq K' \int_{A_0}^{A_t} g(s) ds.$$

The constant K' equals $\ell g(K)$, where ℓ is the sub-multiplicativity constant of g .

3. BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS DRIVEN BY STOCHASTICALLY DISCONTINUOUS MARTINGALES

In this section, we will work on the complete stochastic basis $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ and fix throughout a \mathbb{G} -stopping time T and a (\mathbb{G}, \mathbb{P}) -martingale $X^T \in \mathcal{H}^2$ such that (X^T, C^X) satisfies Assumption (C). Abusing notation, we will refer to the stopped process X^T as simply the process X , since T is given. Hence, the time interval on which we will be working throughout this section will always be the stochastic time interval $\llbracket 0, T \rrbracket$. In addition, a non-decreasing process A will be fixed below, see (3.3). In order to simplify notation, and since there is no danger of confusion, we will omit X from our spaces and norms. Therefore they become, for any $\beta \geq 0$ and $(t, \omega) \in \mathbb{R}_+ \times \Omega$,

$$\mathbb{L}_\beta^2 := \mathbb{L}_\beta^2(\mathcal{G}_T; \mathbb{R}^d), \quad \mathbb{H}_\beta^2 := \mathbb{H}_\beta^2(X), \quad \mathcal{S}_\beta^2 := \mathcal{S}_\beta^2(X), \quad \mathbb{H}_\beta^{2,c} := \mathbb{H}_\beta^2(X^c), \quad \mathbb{H}_\beta^{2,d} := \mathbb{H}_\beta^2(X^d),$$

$$\mathcal{H}_\beta^{2,\perp} := \mathcal{H}_\beta^2(X^\perp), \quad \mathfrak{H}_{t,\omega} := \mathfrak{H}_{t,\omega}^X, \quad \mathfrak{H} := \mathfrak{H}^X,$$

$$\|\cdot\|_\beta := \|\cdot\|_{\beta,X}, \quad \|\cdot\|_{\star,\beta} := \|\cdot\|_{\star,\beta,X}, \quad \|\cdot\|_t := \|\cdot\|_t^X,$$

$$C := C^X, \quad c := c^X, \quad \mu := \mu^X, \quad \nu := \nu^X, \quad \tilde{\mu} := \tilde{\mu}^X$$

When $\beta = 0$, we also suppress it from the notation of the previous spaces.

3.1. Formulation of the problem. We are interested in proving existence and uniqueness of the solution of a backward stochastic differential equation of the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, U_s(\cdot)) dC_s - \int_t^T Z_s^\top dX_s^c - \int_t^T \int_{\mathbb{R}^n} U_s(x) \tilde{\mu}(ds, dx) - \int_t^T dN_s, \quad (3.1)$$

which means that, given the data $(X, \mathbb{G}, T, \xi, f, C)$, we seek a quadruple (Y, Z, U, N) that satisfies equation (3.1), $\mathbb{P} - a.s.$ The martingale X is not assumed quasi-left-continuous and may have stochastic discontinuities. As a result, the process C may also have discontinuities. In other words, we consider BSDEs with jumps that are driven both by continuous-time and by discrete-time martingales in a unified framework.

The data of the BSDE should satisfy the following conditions:

(F1) The martingale X belongs to \mathcal{H}^2 and (X, C) satisfies Assumption (C).

(F2) The terminal condition satisfies $\xi \in \mathbb{L}_{\hat{\beta}}^2$ for some $\hat{\beta} > 0$.

(F3) The generator⁴ of the equation $f : \Omega \times \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^{d \times m} \times \mathfrak{H} \longrightarrow \mathbb{R}^d$ is such that for any $(y, z, u) \in \mathbb{R}^d \times \mathbb{R}^{d \times m} \times \mathfrak{H}$, the map

$$(t, \omega) \longmapsto f(t, \omega, y, z, u_t(\omega; \cdot)) \text{ is } \mathcal{F}_t \otimes \mathcal{B}([0, t])\text{-measurable.}$$

Moreover, f satisfies a stochastic Lipschitz condition, that is to say there exist

$$r : (\Omega \times \mathbb{R}_+, \mathcal{P}) \longrightarrow (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+)) \text{ and } \vartheta = (\theta^c, \theta^d) : (\Omega \times \mathbb{R}_+, \mathcal{P}) \longrightarrow (\mathbb{R}_+^2, \mathcal{B}(\mathbb{R}_+^2)),$$

such that, for $dC_t \otimes d\mathbb{P} - a.e. (t, \omega) \in \mathbb{R}_+ \times \Omega$

$$\begin{aligned} & |f(t, \omega, y, z, u_t(\omega; \cdot)) - f(t, \omega, y', z', u'_t(\omega; \cdot))|^2 \\ & \leq r_t(\omega) |y - y'|^2 + \theta_t^c(\omega) \|c_t(\omega)(z - z')\|^2 + \theta_t^d(\omega) (\|u_t(\omega; \cdot) - u'_t(\omega; \cdot)\|_t(\omega))^2. \end{aligned} \quad (3.2)$$

(F4) Let⁵ $\alpha^2 := \max\{\sqrt{r}, \theta^c, \theta^d\}$ and define the increasing, \mathbb{G} -predictable and càdlàg process

$$A := \int_0^\cdot \alpha_s^2 dC_s. \quad (3.3)$$

Then there exists $\Phi > 0$ such that

$$\Delta A_t(\omega) \leq \Phi, \text{ for } dC_t \otimes d\mathbb{P} - a.e. (t, \omega) \in \mathbb{R}_+ \times \Omega. \quad (3.4)$$

(F5) We have for the same $\hat{\beta}$ as in **(F2)**

$$\mathbb{E} \left[\int_0^T e^{\hat{\beta} A_t} \frac{|f(t, 0, 0, \mathbf{0})|^2}{\alpha_t^2} dC_t \right] < \infty,$$

where $\mathbf{0}$ denotes the null application from \mathbb{R}^n to \mathbb{R} .

Remark 3.1. In the case where the integrator C of the Lebesgue–Stieltjes integral is a continuous process, we can choose between the integrands

$$(f(t, Y_t, Z_t, U_t(\cdot)))_{t \in \llbracket 0, T \rrbracket} \text{ and } (f(t, Y_{t-}, Z_t, U_t(\cdot)))_{t \in \llbracket 0, T \rrbracket},$$

and we still obtain the same solution, as they coincide on a $dC \otimes d\mathbb{P}$ -null set. However, in the case where the integrator C is càdlàg, the corresponding solutions may differ. In the formulation of the problem we have chosen the first one, while all our results can readily be adapted to the second case as well. Nevertheless, in Subsection 3.3 we will see that, in special cases, the conditions for existence and uniqueness of a solution in these two cases can differ significantly.

In classical results on BSDEs, the pair (ξ, f) is called *standard data*. In our case, we generalise the last term and say that the sextuple $(X, \mathbb{G}, T, \xi, f, C)$ is the *standard data under $\hat{\beta}$* , whenever its elements satisfy Assumptions **(F1)–(F5)** for this specific $\hat{\beta}$.

Definition 3.2. A solution of the BSDE (3.1) with standard data $(X, \mathbb{G}, T, \xi, f, C)$ under $\hat{\beta} > 0$ is a quadruple of processes

$$(Y, Z, U, N) \in \mathbb{H}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp} \quad \text{or} \quad (Y, Z, U, N) \in \mathcal{S}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp},$$

for some $\beta \leq \hat{\beta}$ such that, $\mathbb{P} - a.s.$, for any $t \in \llbracket 0, T \rrbracket$,

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, U_s(\cdot)) dC_s - \int_t^T Z_s^\top dX_s^c - \int_t^T \int_{\mathbb{R}^n} U_s(x) \tilde{\mu}(ds, dx) - \int_t^T dN_s. \quad (3.5)$$

⁴This is also called *driver* of the BSDE.

⁵We assume, without loss of generality, that $\alpha_t > 0$, $dC_{t \wedge T} \otimes d\mathbb{P} - a.e.$

Remark 3.3. We emphasise that in (3.5), the stochastic integrals are well defined since $(Z, U, N) \in \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$. Let us verify that the integral

$$\int_0^\cdot f(s, Y_s, Z_s, U_s(\cdot)) dC_s$$

is also well-defined. First of all, we know by definition that for any $(y, z, u) \in \mathbb{R}^d \times \mathbb{R}^{d \times m} \times \mathfrak{H}$, there exists a $dC_t \otimes d\mathbb{P}$ -null set $\mathcal{N}^{y,z,u}$ such that for any $(t, \omega) \notin \mathcal{N}^{y,z,u}$

$$f(t, \omega, y, z, u_t(\omega; \cdot)) \text{ is well defined and } u_t(\omega; \cdot) \in \mathfrak{H}_{t,\omega}.$$

Moreover, by Lemma 2.4, we know also that for some $dC_t \otimes d\mathbb{P}$ -null set $\tilde{\mathcal{N}}$, we have for every $(t, \omega) \notin \tilde{\mathcal{N}}$, that $\mathfrak{H}_{t,\omega}$ is Polish for the norm $\|\cdot\|_t(\omega)$, so that it admits a countable dense subset which we denote by $H_{t,\omega}$. Let us then define

$$H := \left\{ u \in \mathfrak{H}, u_t(\omega; \cdot) \in H_{t,\omega}, \forall (t, \omega) \notin \tilde{\mathcal{N}} \right\}, \quad \mathcal{N} := \bigcup \left\{ \mathcal{N}^{y,z,u}, (y, z, u) \in \mathbb{Q}^d \times \mathbb{Q}^{d \times m} \times H \right\},$$

where \mathbb{Q} and $\mathbb{Q}^{d \times m}$ are the subsets of \mathbb{R} and $\mathbb{R}^{d \times m}$ with rational components.

Then, since H is countable, \mathcal{N} is still a $dC_t \otimes d\mathbb{P}$ -null set. Then, it suffices to use (F3) to realize that for any $(t, \omega) \notin \mathcal{N} \cup \tilde{\mathcal{N}}$, f is continuous in (y, z, u) , and conclude that we can actually define $f(t, \omega, y, z, u_t(\omega; \cdot))$ outside a universal $dC_t \otimes d\mathbb{P}$ -null set. This implies in particular that for any $(Y, Z, U) \in \mathbb{H}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d}$

$$f(t, \omega, Y_t(\omega), Z_t(\omega), U_t(\omega; \cdot)) \text{ is defined for } dC_t \otimes d\mathbb{P} - a.e. (t, \omega) \in \llbracket 0, T \rrbracket \times \Omega.$$

Finally, it suffices to use (F3) and (F5) to conclude that

$$\int_0^T |f(t, \omega, Y_t(\omega), Z_t(\omega), U_t(\omega; \cdot))| dC_t(\omega) \text{ is finite for } dC_t \otimes d\mathbb{P} - a.e. (t, \omega) \in \llbracket 0, T \rrbracket \times \Omega.$$

3.2. Existence and uniqueness: statement. We devote this subsection to the statement of our main theorem. Before that, we need some preliminary results of a purely analytical nature, whose proofs are relegated to Appendix B.

Lemma 3.4. Fix $\beta, \Phi > 0$ and consider the set $\mathcal{C}_\beta := \{(\gamma, \delta) \in (0, \beta]^2, \gamma < \delta\}$. We define the following quantity

$$\Pi^\Phi(\gamma, \delta) := \frac{9}{\delta} + (2 + 9\delta) \frac{e^{(\delta-\gamma)\Phi}}{\gamma(\delta-\gamma)}.$$

Then, the infimum of Π^Φ is given by

$$M^\Phi(\beta) := \inf_{(\gamma, \delta) \in \mathcal{C}_\beta} \Pi^\Phi(\gamma, \delta) = \frac{9}{\beta} + \frac{\Phi^2(2 + 9\beta)}{\sqrt{\beta^2\Phi^2 + 4} - 2} \exp\left(\frac{\beta\Phi + 2 - \sqrt{\beta^2\Phi^2 + 4}}{2}\right),$$

and is attained at the point $(\bar{\gamma}^\Phi(\beta), \beta)$ where

$$\bar{\gamma}^\Phi(\beta) := \frac{\beta\Phi - 2 + \sqrt{4 + \beta^2\Phi^2}}{2\Phi}.$$

In addition, if we define

$$\Pi_\star^\Phi(\gamma, \delta) := \frac{8}{\gamma} + \frac{9}{\delta} + 9\delta \frac{e^{(\delta-\gamma)\Phi}}{\gamma(\delta-\gamma)},$$

then the infimum of Π_\star^Φ is given by $M_\star^\Phi(\beta) := \inf_{(\gamma, \delta) \in \mathcal{C}_\beta} \Pi_\star^\Phi(\gamma, \delta) = \Pi_\star^\Phi(\bar{\gamma}_\star^\Phi(\beta), \beta)$, where $\bar{\gamma}_\star^\Phi(\beta)$ is the unique solution in $(\bar{\gamma}^\Phi(\beta), \beta)$ of the equation with unknown x

$$8(\beta - x)^2 - 9\beta e^{(\beta-x)\Phi}(\Phi x^2 - (\beta\Phi - 2)x - \beta) = 0.$$

Moreover, it holds

$$\lim_{\beta \rightarrow \infty} M^\Phi(\beta) = \lim_{\beta \rightarrow \infty} M_\star^\Phi(\beta) = 9e\Phi.$$

Theorem 3.5. *Let $(X, \mathbb{G}, T, \xi, f, C)$ be standard data under $\hat{\beta}$. If $M^\Phi(\hat{\beta}) < \frac{1}{2}$ (resp. $M_\star^\Phi(\hat{\beta}) < \frac{1}{2}$), then there exists a unique quadruple (Y, Z, U, N) which satisfies (3.5) and with $\|(Y, Z, U, N)\|_{\hat{\beta}} < \infty$ (resp. $\|(Y, Z, U, N)\|_{\star, \hat{\beta}} < \infty$).*

Using the results of Lemma 3.4 and Theorem 3.5, it is immediate that as soon as

$$\Phi < \frac{1}{18e}, \quad (3.6)$$

then there always exists a unique solution of the BSDE for $\hat{\beta}$ large enough.

3.3. Comparison with the related literature.

3.3.1. Some counterexamples. As mentioned already in the introduction, Confortola, Fuhrman and Jacod [43, Section 4.3] provided a counterexample to the existence or the uniqueness of the solution of a BSDE in case the integrator C is not a continuous process. We would like to shed some more light on their counterexample here, and discuss various situations in which a solution may or may not exist.

Let us first rewrite their counterexample using our notation. Let $T > 0$, $r \in (0, T]$, X be a piecewise constant process with potentially a single jump at time r , that is $\mathbb{P}(\Delta X_r \neq 0) = p \in (0, 1)$ and $\mathbb{P}(\{\Delta X_t = 0 \text{ for every } t \in (0, \infty)\}) = 1 - p$. Let $\Pi := \{\omega \in \Omega, \Delta X_r(\omega) \neq 0\}$ and Π^c be its complement. Moreover, let \mathbb{G} be the natural filtration of X , $C = p\mathbb{1}_{[r, \infty)}(\cdot)$, and fix some generator $f : [0, T] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$. Given the structure of the filtration \mathbb{G} , the terminal condition ξ can always have a decomposition of the form

$$\xi(\omega) =: \xi^\Pi \mathbb{1}_\Pi(\omega) + \xi^{\Pi^c} \mathbb{1}_{\Pi^c}(\omega), \quad (\xi^\Pi, \xi^{\Pi^c}) \in \mathbb{R} \times \mathbb{R}.$$

Then, [43] considers the following BSDE

$$Y_t + \int_t^T U_s \mu^X(ds) = \xi + \int_{(t, T]} f(s, Y_{s-}, U_s) dC_s, \quad (3.7)$$

and show that, if the generator has the form $f(t, y, u) = \frac{1}{p}(y + g(u))$ for a deterministic function $g : \mathbb{R} \rightarrow \mathbb{R}$, then the BSDE can admit either infinitely many solutions or none.

Once again because of the structure of \mathbb{G} , one can show that the possible solutions for the BSDE (3.7) necessarily have the following form

$$\begin{aligned} Y_t(\omega) &= Y_0 \mathbb{1}_{[0, r)}(t) + \xi^\Pi \mathbb{1}_{[r, \infty)}(t) \mathbb{1}_\Pi(\omega) + \xi^{\Pi^c} \mathbb{1}_{[r, \infty)}(t) \mathbb{1}_{\Pi^c}(\omega), \\ U_t(\omega) &= v(t) + v^\Pi(t) \mathbb{1}_\Pi(\omega) \mathbb{1}_{(r, T]}(t) + v^{\Pi^c}(t) \mathbb{1}_{\Pi^c}(\omega) \mathbb{1}_{(r, T]}(t), \end{aligned}$$

for some $Y_0 \in \mathbb{R}$ and some deterministic functions $v, v^\Pi, v^{\Pi^c} : [0, T] \rightarrow \mathbb{R}$. However, only the value $v(r)$ is actually involved. By [70, Theorem II.3.26] we have that C is the compensator of X , i.e. the process $\tilde{X} := X - C$ is a \mathbb{G} -martingale. Now we can distinguish between the following cases:

(C1) Consider the BSDE (3.7). Then, there exists a solution if and only if there exists a fixed point, called Y_0^\star , for the equation

$$\xi^\Pi + pf(r, x, \xi^\Pi - \xi^{\Pi^c}) = x.$$

The pair $(Y_0^\star, \xi^\Pi - \xi^{\Pi^c})$ is a solution of (3.7). The solution is unique if and only if the fixed point is unique. In case f is globally Lipschitz with respect to its second argument, i.e.

$$|f(t, y_1, u) - f(t, y_2, u)| \leq L^y |y_1 - y_2|,$$

then, a sufficient condition for the existence and uniqueness of the solution is $L^y \Delta C_r < 1$.

(C2) Consider the following BSDE instead, where the stochastic integral is taken with respect to the compensated jump process

$$Y_t + \int_t^T U_s \tilde{\mu}^X(ds) = \xi + \int_{(t, T]} f(s, Y_{s-}, U_s) dC_s. \quad (3.8)$$

Then, there exists a solution if and only if there exists a fixed point, called Y_0^* , for the equation

$$\xi^\Pi + pf(r, x, \xi^\Pi - \xi^{\Pi^c}) + p(\xi^\Pi - \xi^{\Pi^c}) = x.$$

The pair $(Y_0^*, \xi^\Pi - \xi^{\Pi^c})$ is a solution of (3.8). The solution is unique if and only if the fixed point is unique. In case f is globally Lipschitz with respect to the second argument as above, a sufficient condition for the existence and uniqueness of the solution is again $L^y \Delta C_r < 1$.

- (C3) Consider now a BSDE similar to (3.7), where the integrand of the Lebesgue–Stieltjes integral depends on Y instead of Y_- , *i.e.*

$$Y_t + \int_t^T U_s \mu^X(ds) = \xi + \int_{(t,T]} f(s, Y_s, U_s) dC_s. \quad (3.9)$$

Then, there exists a solution if and only if there exists a fixed point, called $v^*(r)$, for the equation

$$\xi^\Pi - \xi^{\Pi^c} - pf(r, \xi^{\Pi^c}, x) + pf(r, \xi^\Pi, x) = x.$$

The pair $(\xi^{\Pi^c} + pf(r, \xi^{\Pi^c} + pv^*(r), v^*(r)), v^*(r))$ is a solution of (3.9). The solution is unique if and only if the fixed point is unique. In case f is globally Lipschitz with respect to its third argument, *i.e.*

$$|f(t, y, u_1) - f(t, y, u_2)| \leq L^u |u_1 - u_2|,$$

then, a sufficient condition for the existence and uniqueness of the solution is $2L^u \Delta C_r < 1$. This condition is not necessary: let $f'(t, y, u) = \frac{1}{p}(g(y) + u)$, where g is a deterministic function, then it holds that $L^u \Delta C_r = 1$; however, (3.9) admits a unique solution, which is given by the pair

$$(\xi^\Pi + g(\xi^\Pi), \xi^\Pi - \xi^{\Pi^c} + g(\xi^\Pi) - g(\xi^{\Pi^c})).$$

- (C4) Finally, consider a BSDE similar to (3.9) where the stochastic integral is taken with respect to the compensated jump process

$$Y_t + \int_t^T U_s \tilde{\mu}^X(ds) = \xi + \int_{(t,T]} f(s, Y_s, U_s) dC_s. \quad (3.10)$$

Then, there exists a solution if and only if there exists a fixed point, called $v^*(r)$, for the equation

$$\xi^\Pi - \xi^{\Pi^c} - pf(r, \xi^{\Pi^c}, x) + pf(r, \xi^\Pi, x) = x.$$

The pair $(\xi^{\Pi^c} + pf(r, \xi^{\Pi^c}, v^*(r)), v^*(r))$ is a solution of (3.10). The solution is unique if and only if the fixed point is unique. In case f is globally Lipschitz with respect to its third argument as above, a sufficient condition for the existence and uniqueness of the solution is again $2L^u \Delta C_r < 1$. Once again this condition is not necessary; indeed, for f' as in (C3), $L^u \Delta C_r = 1$, while the unique solution of the BSDE (3.10) is the pair

$$((1-p)[\xi^\Pi + g(\xi^\Pi)] + p[\xi^{\Pi^c} + g(\xi^{\Pi^c})], \xi^\Pi - \xi^{\Pi^c} + g(\xi^\Pi) - g(\xi^{\Pi^c})).$$

Now, returning to the original counterexample of [43], we can observe that the sufficient condition $L^y \Delta C_r < 1$ is violated there, which explains why wellposedness issues can arise. However, an important observation here is that the structure of the generator plays a crucial role as well. Indeed, if we consider the same BSDE with the following generator $f(t, y, u) = \ell(y + g(u))$ with $\ell \neq \frac{1}{p}$, then the BSDE admits a unique solution.

Let us finally argue why condition (3.6) rules out this counterexample from our setting. The generator $f(t, y, u) = \frac{1}{p}(y + g(u))$ needs to be Lipschitz so that it fits in our framework, and to satisfy (3.2). Let us further assume that the function g is also Lipschitz, say with associated constant L^g . Then, using Young's Inequality, we can obtain

$$|f(t, y, u) - f(t, y', u')|^2 \leq \frac{1+\varepsilon}{p^2} |y - y'|^2 + \frac{1}{p^2} \left(1 + \frac{(L^g)^2}{\varepsilon}\right) |u - u'|^2, \text{ for every } \varepsilon > 0.$$

Therefore, we have that $\alpha_r^2 = \max \left\{ \sqrt{1+\varepsilon}/p, \left(1 + \frac{(L^g)^2}{\varepsilon}\right)/p^2 \right\}$, and it holds

$$\alpha_r^2 \Delta C_r = \max \left\{ \sqrt{1+\varepsilon}, \left(1 + \frac{(L^g)^2}{\varepsilon}\right)/p \right\} \geq \sqrt{1+\varepsilon} > \frac{1}{18e} \text{ for every } \varepsilon > 0.$$

Remark 3.6. Coming back to Remark 3.1, we observe that the dependence of the integrand on Y or Y_- is not always that innocuous. Indeed, the same BSDE might have a solution in the one formulation but not in the other. Observe furthermore that in the first situation the Lipschitz constant L^y appears in the condition for the existence and uniqueness of a solution, while in the second case the Lipschitz constant L^u appears. As stated already before, in our framework we can treat both cases simultaneously, hence naturally both Lipschitz constants appear in our condition, through the definition of α^2 as the maximum of all the Lipschitz constants.

3.3.2. Related literature. Let us now compare our work with the papers by Bandini [10] and Cohen and Elliott [41] who also consider BSDEs in stochastically discontinuous filtrations. The setting in [41] is rather different from ours. Indeed, in our case a driving martingale X is given right from the start, and as a consequence the process C with respect to which the generator f is integrated is linked to the predictable bracket of X . However, the authors of [41] do not choose any X from the start, but consider instead a general martingale representation theorem involving countably many orthogonal martingales, which only requires the space of square integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ to be separable to hold. Furthermore, their process C can, unlike our case, be chosen arbitrary (in the sense that it does not have to be related to the driving martingales), but with the restriction that it has to be deterministic. Moreover, it has to assign positive measure on every interval, see Definition 5.1 therein, hence C cannot be piecewise constant; the latter would naturally arise from a discrete-time martingale with independent increments, which is exactly the situation one encounters when devising numerical schemes for BSDEs. Therefore, their setting cannot be embedded into our framework, and *vice versa*.

On the other hand, in [10], the author considers a BSDE driven by a pure-jump martingale without an orthogonal component, which is a special case of (3.1). The martingale in this setting should actually have jumps of finite activity, hence many of the interesting models for applications in mathematical finance, such as the generalized hyperbolic, CGMY, and Meixner processes, are excluded. Such a restriction is not present at all in our framework. Otherwise, the assumptions and the conclusion in [10] are analogous to the present work. A direct comparison is however not possible, *i.e.* we cannot deduce the existence and uniqueness results in her work from our setting, since the assumptions are not exactly comparable. In particular, the integrability condition (iii) on page 3 in [10] is not compatible with (F5).

Let us also compare our result with the literature on BSDEs with random terminal time. Royer [105], for instance, considers a BSDE driven by Brownian motion, where the terminal time is a $[0, \infty)$ -valued stopping time. Hence, her setting can be embedded in ours, by assuming the absence of jumps and of the orthogonal component, and further requiring that C is a continuous process. She shows existence and uniqueness of a solution under the assumptions that the generator is uniformly Lipschitz in z and continuous in y , and the terminal condition is bounded. Moreover, she requires that either the generator is strictly monotone in y and $f(t, 0, 0)$ is bounded (for all t) or that the generator is monotone in y and $f(t, 0, 0) = 0$ (for all t). These conditions are not directly covered by our Assumptions (F1)–(F5), however if we consider her conditions and assume in addition that the generator is Lipschitz in y , then we can recover the existence and uniqueness result from our main theorem. Let us point out that BSDEs with constant terminal time are related to semi-linear parabolic PDEs, while BSDEs with random terminal time are associated to semi-linear elliptic PDEs.

We would also like to comment briefly on the choice of the norms we consider here. They are mostly inspired from the ones defined in the seminal work of El Karoui and Huang [51], and are equivalent to the usual norms found in the literature when the process A and time T are both bounded. Bandini [10] uses different spaces, where the norm is defined using the Doléans–Dade stochastic exponential instead of the natural exponential. In our setting where A is allowed to be unbounded, we can only say that our norm dominates hers. This means that we require stronger integrability conditions, but

as a result we will also obtain a solution of the BSDE with stronger integrability properties. In any case, our method could be adapted to this choice of the norm, albeit with modified computations in our estimates. We refer the reader to Remark 3.8 below for a more detailed discussion about the definition of the norms.

Let us conclude this section by commenting on the condition (3.6). We start with the observation that the analysis of the counterexample of Confortola *et al.* [43] made in Sub-sub-section 3.3.1 does not allow for a general statement of wellposedness of the BSDE when $\Phi \geq 1$. In this light, the result of Cohen and Elliott [41, Theorem 6.1], which implies that the condition $\Phi < 1$ ensures the wellposedness of the BSDE, lies in the optimal range for Φ . Analogously in the case of Bandini [10], once her results are translated using the Lipschitz assumption in (F3), $\Phi < \frac{1}{\sqrt{2}}$ also ensures the wellposedness of the BSDE. On the contrary, condition (3.6) which reads as $\Phi < 1/(18e)$, may seem much more restrictive. The first immediate remark we can make is that the stochasticity of the integrator C considerably deteriorates the condition on Φ . In [41] the integrator is deterministic, while in [10] and in our case the integrator is stochastic. However, we would like to remind the reader, that, as explained above, the level of generality we are working with is substantially higher than in these two references. We also want to emphasise the fact that our condition is clearly not the sharpest one possible, but we believe it is the sharpest that can be obtained using our method of proof. The main possibilities for improvement are, in our view, twofold:

- First of all, in specific situations (e.g. T bounded, f Lipschitz, less general driving processes, ...) one could most probably improve the *a priori* estimates of Lemma 3.7 by refining several of the inequalities.
- Second, as highlighted in Remark 3.8, we actually have a degree of freedom in choosing the norms we are interested in. In this paper, we used exponentials, while Bandini [10] used stochastic exponentials, but other choices, leading to potentially better estimates, could also be considered.

We leave this interesting problem of finding the optimal Φ open for future research.

3.4. *A priori* estimates. The method of proof we will use follows and extends the one of El Karoui and Huang [51]. In [51], as well as in Pardoux and Peng [101], the result is obtained using fixed-point arguments and the so-called *a priori* estimates. However, we would like to underline that the proof of such estimates in our case is significantly harder, due to the fact that the process C is not necessarily continuous.

The following result can be seen as the *a priori* estimates for a BSDE whose generator does not depend on the solution. In order to keep notation as simple as possible, as well as to make the link with the data of the problem we consider clearer, we will reuse part of the notation of (F1)–(F5), namely ξ, T, f, C, α and A , only for the next two lemmata.

Lemma 3.7. *Let y be a d -dimensional \mathbb{G} -semimartingale of the form*

$$y_t = \xi + \int_t^T f_s dC_s - \int_t^T d\eta_s, \quad (3.11)$$

where T is a \mathbb{G} -stopping time, $\xi \in \mathbb{L}^2(\mathcal{G}_T; \mathbb{R}^d)$, f is a d -dimensional optional process, C is an increasing, predictable and càdlàg process and $\eta \in \mathcal{H}^2$.

Let $A := \alpha^2 \cdot C$ for some predictable process α . Assume that there exists $\Phi > 0$ such that property (3.4) holds for A . Suppose there exists $\beta \in \mathbb{R}_+$ such that

$$\mathbb{E} \left[e^{\beta A_T} |\xi|^2 \right] + \mathbb{E} \left[\int_0^T e^{\beta A_t} \frac{|f_t|^2}{\alpha_t^2} dC_t \right] < \infty.$$

Then we have for any $(\gamma, \delta) \in (0, \beta]^2$, with $\gamma \neq \delta$,

$$\begin{aligned} \|\alpha y\|_{\mathbb{H}_\delta^2}^2 &\leq \frac{2e^{\delta\Phi}}{\delta} \|\xi\|_{\mathbb{L}_\delta^2}^2 + 2\Lambda^{\gamma, \delta, \Phi} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2, \quad \|y\|_{\mathcal{S}_\delta^2}^2 \leq 8 \|\xi\|_{\mathbb{L}_\delta^2}^2 + \frac{8}{\gamma} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2, \\ \|\eta\|_{\mathcal{H}_\delta^2}^2 &\leq 9 \left(1 + e^{\delta\Phi}\right) \|\xi\|_{\mathbb{L}_\delta^2}^2 + 9 \left(\frac{1}{\gamma \vee \delta} + \delta \Lambda^{\gamma, \delta, \Phi} \right) \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2, \end{aligned}$$

where we have defined

$$\Lambda^{\gamma, \delta, \Phi} := \frac{1 \vee e^{(\delta-\gamma)\Phi}}{\gamma |\delta - \gamma|}.$$

As a consequence, we have

$$\|\alpha y\|_{\mathbb{H}_\delta^2}^2 + \|\eta\|_{\mathcal{H}_\delta^2}^2 \leq \tilde{\Pi}^{\delta, \Phi} \|\xi\|_{\mathbb{L}_\delta^2}^2 + \Pi^\Phi(\gamma, \delta) \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2, \quad (3.12)$$

$$\|y\|_{\mathcal{S}_\delta^2}^2 + \|\eta\|_{\mathcal{H}_\delta^2}^2 \leq \tilde{\Pi}_*^{\delta, \Phi} \|\xi\|_{\mathbb{L}_\delta^2}^2 + \Pi_*^\Phi(\gamma, \delta) \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2, \quad (3.13)$$

where

$$\tilde{\Pi}^{\delta, \Phi} := 9 + \left(9 + \frac{2}{\delta}\right) e^{\delta\Phi} \quad \text{and} \quad \tilde{\Pi}_*^{\delta, \Phi} := 17 + 9e^{\delta\Phi}.$$

Proof. Recall the identity

$$y_t = \xi + \int_t^T f_s dC_s - \int_t^T d\eta_s = \mathbb{E} \left[\xi + \int_t^T f_s dC_s \middle| \mathcal{G}_t \right], \quad (3.14)$$

and introduce the anticipating function

$$F(t) = \int_t^T f_s dC_s. \quad (3.15)$$

For $\gamma \in \mathbb{R}_+$, we have by the Cauchy-Schwarz inequality,

$$\begin{aligned} |F(t)|^2 &\leq \int_t^T e^{-\gamma A_s} dA_s \int_t^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s \leq \int_{A_t}^{A_T} e^{-\gamma A_{L_s}} dA_s \int_t^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s \\ &\leq \int_{A_t}^{A_T} e^{-\gamma s} ds \int_t^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s \leq \frac{1}{\gamma} e^{-\gamma A_t} \int_t^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s, \end{aligned} \quad (3.16)$$

where for the third inequality we used Lemma A.1.(vii). For $t = 0$, since we assumed that

$$\mathbb{E} \left[\int_0^T e^{\beta A_t} \frac{|f_t|^2}{\alpha_t^2} dC_t \right] < \infty,$$

we have that the following holds for $0 < \gamma < \beta$

$$\mathbb{E} \left[|F(0)|^2 \right] < \infty.$$

For $\delta \in \mathbb{R}_+$ and by integrating (3.16) w.r.t. $e^{\delta A_t} dA_t$ it follows

$$\begin{aligned} \int_0^T e^{\delta A_t} |F(t)|^2 dA_t &\stackrel{(3.16)}{\leq} \frac{1}{\gamma} \int_0^T e^{(\delta-\gamma)A_t} \int_t^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s dA_t \\ &= \frac{1}{\gamma} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} \int_0^{s-} e^{(\delta-\gamma)A_t} dA_t dC_s \\ &\leq \frac{1}{\gamma} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} \int_0^s e^{(\delta-\gamma)A_t} dA_t dC_s, \end{aligned} \quad (3.17)$$

where we used Tonelli's Theorem in the equality. We can now distinguish between two cases:

- For $\delta > \gamma$, we apply Corollary 2.6 for $g(x) = e^{(\delta-\gamma)x}$, and inequality (3.17) becomes

$$\begin{aligned} \int_0^T e^{\delta A_t} |F(t)|^2 dA_t &\leq \frac{e^{(\delta-\gamma)\Phi}}{\gamma} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} \int_{A_0}^{A_s} e^{(\delta-\gamma)t} dt dC_s \\ &\leq \frac{e^{(\delta-\gamma)\Phi}}{\gamma(\delta-\gamma)} \int_0^T e^{\delta A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s, \end{aligned} \quad (3.18)$$

which is integrable if $\delta \leq \beta$.

- For $\delta < \gamma$, inequality (3.17) can be rewritten as follows

$$\begin{aligned} \int_0^T e^{\delta A_t} |F(t)|^2 dA_t &\leq \frac{1}{\gamma} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} \int_{A_0}^{A_s} e^{(\delta-\gamma)A_{L_t}} dt dC_s \\ &\stackrel{\text{Lem. A.1.(vii)}}{\leq} \frac{1}{\gamma} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} \int_{A_0}^{A_s} e^{(\delta-\gamma)t} dt dC_s \\ &\leq \frac{1}{\gamma|\delta-\gamma|} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} \left(e^{(\delta-\gamma)A_0} - e^{(\delta-\gamma)A_s} \right) dC_s \\ &\leq \frac{1}{\gamma|\delta-\gamma|} \int_0^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s, \end{aligned} \quad (3.19)$$

which is integrable if $\gamma \leq \beta$.

To sum up, for $\gamma, \delta \in (0, \beta]$, $\gamma \neq \delta$, we have

$$\mathbb{E} \left[\int_0^T e^{\delta A_t} |F(t)|^2 dA_t \right] \leq \Lambda^{\gamma, \delta, \Phi} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2. \quad (3.20)$$

For the estimate of $\|\alpha y\|_{\mathbb{H}_\delta^2}$ we first use the fact that

$$\begin{aligned} \|\alpha y\|_{\mathbb{H}_\delta^2}^2 &= \mathbb{E} \left[\int_0^T e^{\delta A_t} |y_t|^2 dA_t \right] \leq 2\mathbb{E} \left[\int_0^T \mathbb{E} \left[e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(t)|^2 \middle| \mathcal{G}_t \right] dA_t \right] \\ &= 2\mathbb{E} \left[\int_0^\infty \mathbb{E} \left[e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(t)|^2 \middle| \mathcal{G}_t \right] dA_t^T \right] \\ &\stackrel{\text{Cor. C.1}}{=} 2\mathbb{E} \left[\int_0^\infty e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(t)|^2 dA_t^T \right] \\ &= 2\mathbb{E} \left[\int_0^T e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(t)|^2 dA_t \right] \\ &\stackrel{\text{Cor. 2.6}}{\leq} \frac{2e^{\delta\Phi}}{\delta} \|\xi\|_{\mathbb{L}_\delta^2}^2 + 2\Lambda^{\gamma, \delta, \Phi} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2. \end{aligned} \quad (3.21)$$

In the second equality we have used that the processes $|\xi|^2 \mathbf{1}_{\Omega \times [0, \infty]}(\cdot)$ and $|F(\cdot)|^2$ are uniformly integrable, hence their optional projections are well defined. Indeed, using (3.16) and remembering that $\mathbb{E}[|F(0)|^2] < \infty$, we can conclude the uniform integrability of $|F(\cdot)|^2$. Then, by [65, Theorem 5.4] it holds that

$$\begin{aligned} o \left(e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(\cdot)|^2 \right)_t &= e^{\delta A_t} o \left(|\xi|^2 \mathbf{1}_{\Omega \times [0, \infty]}(\cdot) \right)_t + e^{\delta A_t} o \left(|F(\cdot)|^2 \right)_t \\ &= e^{\delta A_t} \mathbb{E} \left[|\xi|^2 \middle| \mathcal{G}_t \right] + e^{\delta A_t} \mathbb{E} \left[|F(t)|^2 \middle| \mathcal{G}_t \right] \\ &= \mathbb{E} \left[e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(t)|^2 \middle| \mathcal{G}_t \right], \end{aligned}$$

which justifies the use of Corollary C.1.

For the estimate of $\|y\|_{\mathcal{S}_\delta^2}$ we have

$$\begin{aligned}
\|y\|_{\mathcal{S}_\delta^2} &= \mathbb{E} \left[\sup_{0 \leq t \leq T} \left(e^{\frac{\delta}{2} A_t} |y_t| \right)^2 \right] \\
&\leq 2 \mathbb{E} \left[\sup_{0 \leq t \leq T} \mathbb{E} \left[\sqrt{e^{\delta A_t} |\xi|^2 + e^{\delta A_t} |F(t)|^2} \middle| \mathcal{G}_t \right]^2 \right] \\
&\leq 2 \mathbb{E} \left[\sup_{0 \leq t \leq T} \mathbb{E} \left[\sqrt{e^{\delta A_T} |\xi|^2 + \frac{1}{\gamma} e^{(\delta-\gamma)A_t} \int_t^T e^{\gamma A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s} \middle| \mathcal{G}_t \right]^2 \right] \\
&\leq 2 \mathbb{E} \left[\sup_{0 \leq t \leq T} \mathbb{E} \left[\sqrt{e^{\delta A_T} |\xi|^2 + \frac{1}{\gamma} \int_0^T e^{(\gamma+(\delta-\gamma)^+)A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s} \middle| \mathcal{G}_t \right]^2 \right] \\
&\leq 8 \mathbb{E} \left[e^{\delta A_T} |\xi|^2 + \frac{1}{\gamma} \int_0^T e^{(\gamma \vee \delta) A_s} \frac{|f_s|^2}{\alpha_s^2} dC_s \right] \leq 8 \|\xi\|_{\mathbb{L}_\delta^2}^2 + \frac{8}{\gamma} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2 \tag{3.22}
\end{aligned}$$

for $\gamma \vee \delta \leq \beta$, where in the second and third inequalities we used the inequality $a + b \leq \sqrt{2(a^2 + b^2)}$ and (3.16) respectively.

What remains is to control $\|\eta\|_{\mathcal{H}_\delta^2}$. We remind once more the reader that $\int_t^T d\eta_s = \xi - y_t + F(t)$, hence

$$\mathbb{E} \left[|\xi - y_t - F(t)|^2 \middle| \mathcal{G}_t \right] = \mathbb{E} \left[\int_t^T d\text{Tr}[\langle \eta \rangle_s] \middle| \mathcal{G}_t \right]. \tag{3.23}$$

In addition, we have

$$\begin{aligned}
\int_0^T e^{\delta A_s} d\text{Tr}[\langle \eta \rangle_s] &= \int_0^T \int_{A_0}^{A_s} \delta e^{\delta t} dt d\text{Tr}[\langle \eta \rangle_s] + \text{Tr}[\langle \eta \rangle_T] \\
&\stackrel{\text{Lem. A.1. (vii)}}{\leq} \int_0^T \int_{A_0}^{A_s} \delta e^{\delta A_{L_t}} dt d\text{Tr}[\langle \eta \rangle_s] + \text{Tr}[\langle \eta \rangle_T] \\
&= \delta \int_0^T \int_0^s e^{\delta A_t} dA_t d\text{Tr}[\langle \eta \rangle_s] + \text{Tr}[\langle \eta \rangle_T] \\
&\leq \delta \int_0^T e^{\delta A_t} \int_t^T d\text{Tr}[\langle \eta \rangle_s] dA_t + \text{Tr}[\langle \eta \rangle_T],
\end{aligned}$$

so that

$$\|\eta\|_{\mathcal{H}_\delta^2} \leq \delta \mathbb{E} \left[\int_0^T e^{\delta A_t} \int_t^T d\text{Tr}[\langle \eta \rangle_s] dA_t \right] + \mathbb{E} [\text{Tr}[\langle \eta \rangle_T]]. \tag{3.24}$$

For the first summand on the right-hand-side of (3.24), we have

$$\begin{aligned}
\mathbb{E} \left[\int_0^T e^{\delta A_t} \int_t^T d\text{Tr}[\langle \eta \rangle_s] dA_t \right] &\stackrel{\text{Cor. 3.1}}{=} \mathbb{E} \left[\int_0^T e^{\delta A_t} \mathbb{E} \left[\int_t^T d\text{Tr}[\langle \eta \rangle_s] \middle| \mathcal{G}_t \right] dA_t \right] \\
&\stackrel{(3.23)}{=} \mathbb{E} \left[\int_0^T e^{\delta A_t} \mathbb{E} \left[|\xi - y_t + F(t)|^2 \middle| \mathcal{G}_t \right] dA_t \right] \\
&\leq 3 \mathbb{E} \left[\int_0^T e^{\delta A_t} \mathbb{E} \left[|\xi|^2 + |y_t|^2 + |F(t)|^2 \middle| \mathcal{G}_t \right] dA_t \right] \\
&\stackrel{(3.14)}{\leq} 3 \mathbb{E} \left[\int_0^T e^{\delta A_t} |\xi|^2 dA_t \right] + 3 \mathbb{E} \left[\int_0^T e^{\delta A_t} |F(t)|^2 dA_t \right] \\
&\quad + 6 \mathbb{E} \left[\int_0^T e^{\delta A_t} \mathbb{E} \left[|\xi|^2 + |F(t)|^2 \middle| \mathcal{G}_t \right] dA_t \right] \\
&\stackrel{\text{Cor. 3.1}}{=} 9 \mathbb{E} \left[\int_0^T e^{\delta A_t} |\xi|^2 dA_t \right] + 9 \mathbb{E} \left[\int_0^T e^{\delta A_t} |F(t)|^2 dA_t \right] \\
&\stackrel{\text{Cor. 2.6}}{\leq} \frac{9e^{\delta\Phi}}{(3.20)} \|\xi\|_{\mathbb{L}_\delta^2}^2 + 9\Lambda^{\gamma, \delta, \Phi} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2.
\end{aligned}$$

We now need an estimate for $\mathbb{E} \left[\int_0^T d\text{Tr}[\langle \eta \rangle_s] \right]$, i.e. the second summand of (3.24), which is given by

$$\begin{aligned}
\mathbb{E} [\text{Tr}[\langle \eta \rangle_T]] &= \mathbb{E} [|\xi - y_0 + F(0)|^2] \leq 3 \mathbb{E} [|\xi|^2 + |y_0|^2 + |F(0)|^2] \\
&\stackrel{(3.14)}{\leq} 9 \mathbb{E} [|\xi|^2] + 9 \mathbb{E} [|F(0)|^2] \stackrel{(3.16)}{\leq} 9 \|\xi\|_{\mathbb{L}_\delta^2}^2 + \frac{9}{\gamma \vee \delta} \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2,
\end{aligned}$$

where we used the fact that $\mathbb{E} [|y_0|^2] \leq 2 \mathbb{E} [|\xi|^2 + |F(0)|^2]$.

Then (3.24) yields

$$\|\eta\|_{\mathcal{H}_\delta^2}^2 \leq 9 \left(1 + e^{\delta\Phi} \right) \|\xi\|_{\mathbb{L}_\delta^2}^2 + 9 \left(\frac{1}{\gamma \vee \delta} + \delta \Lambda^{\gamma, \delta, \Phi} \right) \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_{\gamma \vee \delta}^2}^2. \quad (3.25)$$

Remark 3.8. An alternative framework can be provided if we define the norms in Subsection 2.3 using another positive and increasing function h instead of the exponential function. In order to obtain the required *a priori* estimates, we need to assume that h is sub-multiplicative⁶ and that it shares some common properties with the exponential function. The following provides the *a priori* estimates of the semi-martingale decomposition (3.11) in the case $h : \mathbb{R} \rightarrow [1, \infty)$, with $h(x) = (1+x)^\zeta$, for $\zeta \geq 1$, with the additional assumption that the process A defined in (F4) is \mathbb{P} -a.s. bounded by Ψ . It holds for $\frac{1}{\zeta} < \gamma < \delta < \hat{\beta}$

$$\begin{aligned}
\|\alpha y\|_{\mathbb{H}_\delta^2}^2 + \|\eta\|_{\mathbb{H}_\delta^2}^2 &\leq \left[2h(\Psi)h^\delta(\Phi) + 9 + \frac{9[h(\Psi)]^{1-\frac{1}{\zeta}}[h(\Phi)]^{\delta-\frac{1}{\zeta}}}{\delta - \frac{1}{\zeta} + 1} \right] \|\xi\|_{\mathbb{L}_\delta^2}^2 \\
&\quad + \left[\frac{2[h(\Psi)]^{1+\frac{1}{\zeta}}[h(\Phi)]^{\delta-\gamma+\frac{1}{\zeta}}}{\delta - \gamma + \frac{1}{\zeta} + 1} + \frac{9h(\Psi)[h(\Phi)]^{\delta-\gamma}}{\delta - \gamma + 1} + \frac{9}{\delta\zeta - 1} \right] \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_\delta^2}^2, \\
\|y\|_{\mathcal{S}_\delta^2}^2 + \|\eta\|_{\mathbb{H}_\delta^2}^2 &\leq \left[8 + 2h(\Psi)h^\delta(\Phi) \right] \|\xi\|_{\mathbb{H}_\delta^2}^2 \\
&\quad + \left[\frac{[h(\Psi)]^{\frac{1}{\zeta}}}{\gamma\zeta - 1} + \frac{2[h(\Psi)]^{1+\frac{1}{\zeta}}[h(\Phi)]^{\delta-\gamma+\frac{1}{\zeta}}}{\delta - \gamma + \frac{1}{\zeta} + 1} \right] \left\| \frac{f}{\alpha} \right\|_{\mathbb{H}_\delta^2}^2.
\end{aligned}$$

⁶In the proof of [107, Proposition 25.4] we can find a convenient tool for constructing sub-multiplicative functions.

Let us also provide the following pathwise estimates.

Lemma 3.9. *For the \mathbb{G} -semimartingale y with decomposition (3.11), we have*

$$\int_0^T |y_t|^2 dA_t \leq 2A_T \left(\sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + \sup_{t \in [0, T]} \left| \mathbb{E} \left[\int_0^T |f_s| dC_s \middle| \mathcal{G}_t \right] \right|^2 \right).$$

Moreover, for the martingale part $\eta \in \mathcal{H}^2$, we have

$$\sup_{t \in [0, T]} |\eta_t - \eta_0|^2 \leq 12 \sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + 12 \sup_{t \in [0, T]} \left| \mathbb{E} \left[\int_0^T |f_s| dC_s \middle| \mathcal{G}_t \right] \right|^2 + 3 \left| \int_0^T |f_s| dC_s \right|^2.$$

Proof. By Identity (3.14), we have

$$|y_t|^2 = \left| \mathbb{E} \left[\xi + \int_t^T f_s dC_s \middle| \mathcal{G}_t \right] \right|^2 \leq 2 |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + 2 \left| \mathbb{E} \left[\int_t^T f_s dC_s \middle| \mathcal{G}_t \right] \right|^2, \quad (3.26)$$

which implies

$$\begin{aligned} \int_0^T |y_t|^2 dA_t &\leq 2A_T \sup_{t \in [0, T]} \left\{ |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + \left| \mathbb{E} \left[\int_t^T f_s dC_s \middle| \mathcal{G}_t \right] \right|^2 \right\} \\ &\leq 2A_T \left(\sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + \sup_{t \in [0, T]} \left(\sum_{i=1}^d \mathbb{E} \left[\int_0^T |f_s^i| dC_s \middle| \mathcal{G}_t \right] \right)^2 \right) \\ &\leq 2A_T \left(\sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + \sup_{t \in [0, T]} \left| \mathbb{E} \left[\int_0^T |f_s| dC_s \middle| \mathcal{G}_t \right] \right|^2 \right). \end{aligned}$$

For the martingale part, by Identity (3.14), we have $\eta_t - \eta_0 = y_t - y_0 + \int_0^t f_s dC_s$ which leads to

$$\begin{aligned} \sup_{t \in [0, T]} |\eta_t - \eta_0|^2 &\leq 3|y_t|^2 + 3|y_0|^2 + 3 \left| \int_0^t f_s dC_s \right|^2 \\ &\stackrel{(3.26)}{\leq} 12 \sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + 12 \sup_{t \in [0, T]} \left| \mathbb{E} \left[\int_t^T f_s dC_s \middle| \mathcal{G}_t \right] \right|^2 + 3 \left| \int_0^t f_s dC_s \right|^2 \\ &\stackrel{(3.26)}{\leq} 12 \sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + 12 \sup_{t \in [0, T]} \left(\sum_{i=1}^d \mathbb{E} \left[\int_t^T |f_s^i| dC_s \middle| \mathcal{G}_t \right] \right)^2 + 3 \sum_{i=1}^d \left(\int_0^t |f_s^i| dC_s \right)^2 \\ &\leq 12 \sup_{t \in [0, T]} |\mathbb{E}[\xi | \mathcal{G}_t]|^2 + 12 \sup_{t \in [0, T]} \left| \mathbb{E} \left[\int_0^T |f_s| dC_s \middle| \mathcal{G}_t \right] \right|^2 + 3 \left| \int_0^T |f_s| dC_s \right|^2. \quad \square \end{aligned}$$

Remark 3.10. Viewing (3.11) as a BSDE whose generator does not depend on y and η , then this BSDE has a solution, which can be uniquely determined by the pair (y, η) . Indeed, consider the data $(\mathbb{G}, T, \xi, f, C)$ and the processes α and A , which all satisfy the respective assumptions of Lemma 3.7 for some $\hat{\beta} > 0$. Then the semimartingale

$$y_t = \mathbb{E} \left[\xi + \int_t^T f_s dC_s \middle| \mathcal{G}_t \right] = \mathbb{E} \left[\xi + \int_0^T f_s dC_s \middle| \mathcal{G}_t \right] - \int_0^t f_s dC_s, \quad t \in \mathbb{R}_+$$

satisfies $y_T = \xi$ and for $\eta := \mathbb{E}[\xi + \int_0^T f_s dC_s | \mathcal{G}]$

$$\begin{aligned} y_t - y_T &= \mathbb{E} \left[\xi + \int_t^T f_s dC_s \middle| \mathcal{G}_t \right] - \xi = \mathbb{E} \left[\xi + \int_0^T f_s dC_s \middle| \mathcal{G}_t \right] - \int_0^t f_s dC_s - \xi \\ &= \mathbb{E} \left[\xi + \int_0^T f_s dC_s \middle| \mathcal{G}_t \right] + \int_t^T f_s dC_s - \int_0^T f_s dC_s - \xi = \eta_t + \int_t^T f_s dC_s - \eta_T. \end{aligned}$$

Now, one possible choice of a square-integrable \mathbb{G} -martingale X that we can choose such that $(X, \mathbb{G}, T, \xi, f, C)$ become standard data for any arbitrarily chosen integrator C , is the zero martingale. Hence, given the standard data $(0, \mathbb{G}, T, \xi, f, C)$ the quadruple (y, Z, U, η) satisfies the BSDE

$$y_t = \xi + \int_t^T f_s dC_s - \int_t^T d\eta_s, \quad t \in \llbracket 0, T \rrbracket,$$

for any pair (Z, U) . Assume now that there exists a quadruple $(\tilde{y}, \tilde{Z}, \tilde{U}, \tilde{\eta})$ which satisfies

$$\tilde{y}_t = \xi + \int_t^T f_s dC_s - \int_t^T d\tilde{\eta}_s, \quad t \in \llbracket 0, T \rrbracket.$$

Then, the pair $(y - \tilde{y}, \eta - \tilde{\eta})$ satisfies

$$y - \tilde{y}_t = - \int_t^T d(\eta - \tilde{\eta})_s, \quad t \in \llbracket 0, T \rrbracket,$$

and by Lemma 3.7, for $\xi = 0$ and $f = 0$, we conclude that $\|y - \tilde{y}\|_{\mathcal{S}^2} = \|\eta - \tilde{\eta}\|_{\mathcal{H}^2} = 0$. Therefore y and \tilde{y} , resp. η and $\tilde{\eta}$, are indistinguishable, which implies our initial statement that every solution can be uniquely determined by the pair (y, η) .

In order to obtain the *a priori* estimates for the BSDE (3.1), we will have to consider solutions (Y^i, Z^i, U^i, N^i) , $i = 1, 2$, associated with the data $(X, \mathbb{G}, T, \xi^i, f^i, C)$, $i = 1, 2$ under $\hat{\beta}$, where we also assume that f^1, f^2 have common r, ϑ bounds. Denote the difference between the two solutions by $(\delta Y, \delta Z, \delta U, \delta N)$, as well as $\delta \xi := \xi^1 - \xi^2$ and

$$\delta_2 f_t := (f^1 - f^2)(t, Y_t^2, Z_t^2, U_t^2(\cdot)), \quad \psi_t := f^1(t, Y_t^1, Z_t^1, U_t^1(\cdot)) - f^2(t, Y_t^2, Z_t^2, U_t^2(\cdot)).$$

We have the identity

$$\delta Y_t = \delta \xi + \int_t^T \psi_s dC_s - \int_t^T \delta Z_s dX_s^c - \int_t^T \int_{\mathbb{R}^n} \delta U_s(x) \tilde{\mu}(ds, dx) - \int_t^T d\delta N_s. \quad (3.27)$$

For the wellposedness of this last BSDE we need the following lemma.

Lemma 3.11. *The processes*

$$\int_0^\cdot \delta Z_s dX_s^c \quad \text{and} \quad \int_0^\cdot \int_{\mathbb{R}^n} \delta U_s(x) \tilde{\mu}(dt, dx)$$

are square-integrable martingales with finite associated $\|\cdot\|_{\hat{\beta}}$ -norms.

Proof. The square-integrability is obvious. The inequalities

$$\begin{aligned} \mathbb{E} [\text{Tr}[\langle \delta Z \cdot X^c \rangle]] &\leq 2\mathbb{E} [\text{Tr}[\langle Z^1 \cdot X^c \rangle]] + 2\mathbb{E} [\text{Tr}[\langle Z^2 \cdot X^c \rangle]], \\ \mathbb{E} [\text{Tr}[\langle \delta U \star \tilde{\mu} \rangle]] &\leq 2\mathbb{E} [\text{Tr}[\langle U^1 \star \tilde{\mu} \rangle]] + 2\mathbb{E} [\text{Tr}[\langle U^2 \star \tilde{\mu} \rangle]], \end{aligned}$$

together with Lemma 2.3 guarantee that

$$\mathbb{E} \left[\int_0^T e^{\hat{\beta} A_t} |c_t \delta Z_t|^2 dC_t \right] + \mathbb{E} \left[\int_0^T e^{\hat{\beta} A_t} \|\delta U\|_t^2 dC_t \right] < \infty. \quad \square$$

Therefore, by defining

$$H_t := \int_0^t \delta Z_s dX_s^c + \int_0^t \int_{\mathbb{R}^n} \delta U_s(x) \tilde{\mu}(dt, dx) + \int_0^t d\delta N_s, \quad (3.28)$$

we can treat the BSDE (3.27) exactly as the BSDE (3.11), where the martingale H will play the role of the martingale η .

Proposition 3.12 (*A priori estimates for the BSDE (3.1)*). *Let $(X, \mathbb{G}, T, \xi^i, f^i, C)$, be standard data under $\hat{\beta}$ for $i = 1, 2$. Then $\psi/\alpha \in \mathbb{H}_{\hat{\beta}}^2$ and, if $M^\Phi(\hat{\beta}) < 1/2$, the following estimates hold*

$$\begin{aligned} \|(\alpha\delta Y, \delta Z, \delta U, \delta N)\|_{\hat{\beta}}^2 &\leq \tilde{\Sigma}^\Phi(\hat{\beta}) \|\delta\xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \Sigma^\Phi(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2, \\ \|(\delta Y, \delta Z, \delta U, \delta N)\|_{\star, \hat{\beta}}^2 &\leq \tilde{\Sigma}_\star^\Phi(\hat{\beta}) \|\delta\xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \Sigma_\star^\Phi(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2, \end{aligned}$$

where

$$\begin{aligned} \tilde{\Sigma}^\Phi(\hat{\beta}) &:= \frac{\tilde{\Pi}^{\hat{\beta}, \Phi}}{1 - 2M^\Phi(\hat{\beta})}, \quad \tilde{\Sigma}_\star^\Phi(\hat{\beta}) := \min \left\{ \tilde{\Pi}_\star^{\hat{\beta}, \Phi} + 2M_\star^\Phi(\hat{\beta}), \tilde{\Sigma}^\Phi(\hat{\beta}), 8 + \frac{16}{\hat{\beta}} \tilde{\Sigma}^\Phi(\hat{\beta}) \right\}, \\ \Sigma^\Phi(\hat{\beta}) &:= \frac{2M^\Phi(\hat{\beta})}{1 - 2M^\Phi(\hat{\beta})}, \quad \Sigma_\star^\Phi(\hat{\beta}) := \min \left\{ 2M_\star^\Phi(\hat{\beta})(1 + \Sigma^\Phi(\hat{\beta})), \frac{16}{\hat{\beta}}(1 + \Sigma^\Phi(\hat{\beta})) \right\}. \end{aligned}$$

Proof. For the integrability of ψ , using the Lipschitz property (F3) of f^1, f^2 , we get

$$|\psi_t|^2 \leq 2r_t |\delta Y_t|^2 + 2\theta_t^c \|c_t \delta Z_t\|^2 + 2\theta_t^d \|\delta U_t\|_t^2 + 2|\delta_2 f_t|^2.$$

Hence by the definition of α , which implies that $\frac{r}{\alpha^2} \leq \alpha^2$ and the obvious $\frac{\theta^c}{\alpha^2}, \frac{\theta^d}{\alpha^2} \leq 1$, we get

$$\begin{aligned} \frac{|\psi_t|^2}{\alpha_t^2} &\leq 2 \left(\alpha_t^2 |\delta Y_t|^2 + \|c_t \delta Z_t\|^2 + \|\delta U_t(\cdot)\|_t^2 + \frac{|\delta_2 f|^2}{\alpha^2} \right) \\ &\leq 2\alpha_t^2 |\delta Y_t|^2 + 2\|c_t \delta Z_t\|^2 + 2\|\delta U_t(\cdot)\|_t^2 \\ &\quad + \frac{4}{\alpha^2} \left(|f^1(s, 0, 0, \mathbf{0})|^2 + r_t |Y_t^2|^2 + \theta_t^c \|c_t Z_t^2\|^2 + \theta_t^d \|\delta U_t^2(\cdot)\|_t^2 \right) \\ &\quad + \frac{4}{\alpha^2} \left(|f^2(s, 0, 0, \mathbf{0})|^2 + r_t |\delta Y_t|^2 + \theta_t^c \|c_t \delta Z_t^2\|^2 + \theta_t^d \|\delta U_t(\cdot)\|_t^2 \right) \\ &\leq 6(\alpha_t^2 |\delta Y_t|^2 + \|c_t \delta Z_t\|^2 + \|\delta U_t(\cdot)\|_t^2) + \frac{4}{\alpha^2} \left(|f^1(s, 0, 0, \mathbf{0})|^2 + |f^2(s, 0, 0, \mathbf{0})|^2 \right), \end{aligned} \tag{3.29}$$

where, having used once more that $\frac{r}{\alpha^2} \leq \alpha^2$ and $\frac{\theta^c}{\alpha^2}, \frac{\theta^d}{\alpha^2} \leq 1$, it follows that $\frac{\psi}{\alpha} \in \mathbb{H}_{\hat{\beta}}^2$. Next, for the $\|\cdot\|_{\hat{\beta}}$ -norm, we have

$$\begin{aligned} \|(\delta Y, \delta Z, \delta U, \delta N)\|_{\hat{\beta}}^2 &= \|\alpha\delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta U\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 + \|\delta N\|_{\mathcal{H}_{\hat{\beta}}^{2,\perp}}^2 \\ &\stackrel{(3.28)}{=} \|\alpha\delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|H\|_{\mathcal{H}_{\hat{\beta}}^2}^2 \\ &\stackrel{(3.12)}{\leq} \tilde{\Pi}^{\hat{\beta}, \Phi} \|\delta\xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + M^\Phi(\hat{\beta}) \left\| \frac{\psi}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \\ &\leq \tilde{\Pi}^{\hat{\beta}, \Phi} \|\delta\xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + 2M^\Phi(\hat{\beta}) \left(\|\alpha\delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta U\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 \right) \\ &\quad + 2M^\Phi(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \\ &\leq \tilde{\Pi}^{\hat{\beta}, \Phi} \|\delta\xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + 2M^\Phi(\hat{\beta}) \left(\|\alpha\delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|H\|_{\mathcal{H}_{\hat{\beta}}^2}^2 \right) + 2M^\Phi(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2. \end{aligned}$$

Therefore, this implies

$$\|(\alpha\delta Y, \delta Z, \delta U, \delta N)\|_{\hat{\beta}}^2 \leq \tilde{\Sigma}^\Phi(\hat{\beta}) \|\delta\xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \Sigma^\Phi(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2. \tag{3.30}$$

We can obtain *a priori* estimates for the $\|\cdot\|_{\star,\hat{\beta}}$ -norm by arguing in two different ways:

- The identity (3.27) gives

$$\begin{aligned}
\|(\delta Y, \delta Z, \delta U, \delta N)\|_{\star,\hat{\beta}}^2 &= \|\delta Y\|_{\mathcal{S}_{\hat{\beta}}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta U\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 + \|\delta N\|_{\mathcal{H}_{\hat{\beta}}^{2,\perp}}^2 \\
&\stackrel{(3.28)}{=} \|\delta Y\|_{\mathcal{S}_{\hat{\beta}}^2}^2 + \|H\|_{\mathcal{H}_{\hat{\beta}}^2}^2 \stackrel{(3.13)}{\leq} \tilde{\Pi}_{\star}^{\hat{\beta},\Phi} \|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + M_{\star}^{\Phi}(\hat{\beta}) \left\| \frac{\psi}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \\
&\stackrel{(3.29)}{\leq} \tilde{\Pi}_{\star}^{\hat{\beta},\Phi} \|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + 2M_{\star}^{\Phi}(\hat{\beta}) \|\alpha \delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + 2M_{\star}^{\Phi}(\hat{\beta}) \|H\|_{\mathcal{H}_{\hat{\beta}}^2}^2 + 2M_{\star}^{\Phi}(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \\
&\stackrel{(3.30)}{\leq} \tilde{\Pi}_{\star}^{\hat{\beta},\Phi} \|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + 2M_{\star}^{\Phi}(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \\
&\quad + 2M_{\star}^{\Phi}(\hat{\beta}) \left(\tilde{\Sigma}^{\Phi}(\hat{\beta}) \|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \Sigma^{\Phi}(\hat{\beta}) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \right) \\
&= \left(\tilde{\Pi}_{\star}^{\hat{\beta},\Phi} + 2M_{\star}^{\Phi}(\hat{\beta}) \tilde{\Sigma}^{\Phi}(\hat{\beta}) \right) \|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + 2M_{\star}^{\Phi}(\hat{\beta}) \left(1 + \Sigma^{\Phi}(\hat{\beta}) \right) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2.
\end{aligned}$$

- The identity (3.14) gives

$$\begin{aligned}
\|\delta Y\|_{\mathcal{S}_{\hat{\beta}}^2}^2 &= \mathbb{E} \left[\sup_{0 \leq t \leq T} \left(e^{\frac{\hat{\beta}}{2} A_t} |\delta Y_t| \right)^2 \right] \stackrel{(3.14)}{\leq} \mathbb{E} \left[\sup_{0 \leq t \leq T} \mathbb{E} \left[e^{\frac{\hat{\beta}}{2} A_t} |\delta \xi| + e^{\frac{\hat{\beta}}{2} A_t} \left| \int_t^T \psi_s dC_s \right| \middle| \mathcal{G}_t \right]^2 \right] \\
&\stackrel{(3.16)}{\leq} 2\mathbb{E} \left[\sup_{0 \leq t \leq T} \mathbb{E} \left[\sqrt{e^{\hat{\beta} A_t} |\delta \xi|^2 + \frac{1}{\hat{\beta}} \int_t^T e^{\hat{\beta} A_s} \frac{|\psi_s|^2}{\alpha_s^2} dC_s} \middle| \mathcal{G}_t \right]^2 \right] \\
&\leq 2\mathbb{E} \left[\sup_{0 \leq t \leq T} \mathbb{E} \left[\sqrt{e^{\hat{\beta} A_T} |\delta \xi|^2 + \frac{1}{\hat{\beta}} \int_0^T e^{\hat{\beta} A_s} \frac{|\psi_s|^2}{\alpha_s^2} dC_s} \middle| \mathcal{G}_t \right]^2 \right] \\
&\leq 8\mathbb{E} \left[e^{\hat{\beta} A_T} |\delta \xi|^2 + \frac{1}{\hat{\beta}} \int_0^T e^{\hat{\beta} A_s} \frac{|\psi_s|^2}{\alpha_s^2} dC_s \right] \\
&\stackrel{(3.29)}{\leq} 8\|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \frac{16}{\hat{\beta}} \left(\left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|\alpha \delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta U\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 \right), \tag{3.31}
\end{aligned}$$

where, in the second and fifth inequality we used the inequality $a + b \leq \sqrt{2(a^2 + b^2)}$ and Doob's inequality respectively. Then we can derive the required estimate

$$\begin{aligned}
\|(\delta Y, \delta Z, \delta U, \delta N)\|_{\star,\hat{\beta}}^2 &= \|\delta Y\|_{\mathcal{S}_{\hat{\beta}}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta U\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 + \|\delta N\|_{\mathcal{H}_{\hat{\beta}}^{2,\perp}}^2 \\
&\stackrel{(3.28)}{=} \|\delta Y\|_{\mathcal{S}_{\hat{\beta}}^2}^2 + \|H\|_{\mathcal{H}_{\hat{\beta}}^2}^2 \\
&\stackrel{(3.31)}{\leq} 8\|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \frac{16}{\hat{\beta}} \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \frac{16}{\hat{\beta}} \|\alpha \delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \frac{16}{\hat{\beta}} \|H\|_{\mathcal{H}_{\hat{\beta}}^2}^2 \\
&\stackrel{(3.28)}{\leq} \left(8 + \frac{16}{\hat{\beta}} \tilde{\Sigma}^{\Phi}(\hat{\beta}) \right) \|\delta \xi\|_{\mathbb{L}_{\hat{\beta}}^2}^2 + \frac{16}{\hat{\beta}} \left(1 + \Sigma^{\Phi}(\hat{\beta}) \right) \left\| \frac{\delta_2 f}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2. \quad \square
\end{aligned}$$

3.5. Proof of the main theorem. We will use now the previous estimates to obtain the existence of a unique solution using a fixed point argument.

Proof of Theorem 3.5. Let (y, z, u, n) be such that $(\alpha y, z, u, n) \in \mathbb{H}_{\hat{\beta}}^2 \times \mathbb{H}_{\hat{\beta}}^{2,c} \times \mathbb{H}_{\hat{\beta}}^{2,d} \times \mathcal{H}^{2,\perp}$. Then the process M defined by

$$M := \mathbb{E} \left[\xi + \int_0^T f(s, y_s, z_s, u_s(\cdot)) dC_s \middle| \mathcal{G} \right] + n. \in \mathcal{H}^2,$$

and by [70, Lemma III.4.24] it has a unique, up to indistinguishability, orthogonal decomposition

$$M = M_0 + \int_0^\cdot Z_s dX_s^c + \int_0^\cdot \int_{\mathbb{R}^n} U_s(x) \tilde{\mu}(ds, dx) + L.,$$

where $(Z, U, L) \in \mathbb{H}^{2,c} \times \mathbb{H}^{2,d} \times \mathcal{H}^{2,\perp}$. In view of the identity

$$M_T - M_t = \int_t^T Z_s dX_s^c + \int_t^T \int_{\mathbb{R}^n} U_s(x) \tilde{\mu}(ds, dx) + \int_t^T dL_s, \quad 0 \leq t \leq T,$$

we obtain

$$\begin{aligned} \mathbb{E} \left[\xi + \int_t^T f(s, y_s, z_s, u_s(\cdot)) dC_s \middle| \mathcal{G}_t \right] &= \xi + \int_t^T f(s, y_s, z_s, u_s(\cdot)) dC_s - \int_t^T Z_s dX_s^c \\ &\quad - \int_t^T \int_{\mathbb{R}^n} U_s(x) \tilde{\mu}(ds, dx) - \int_t^T dN_s, \end{aligned}$$

where $N := L - n$. Define

$$Y_t := \mathbb{E} \left[\xi + \int_t^T f(s, y_s, z_s, u_s(\cdot)) dC_s \middle| \mathcal{G}_t \right].$$

In order to construct a contraction using Lemma 3.7, we need to choose $\delta > \gamma$. Then by Lemma 3.4 we can choose $\gamma^* \in (0, \hat{\beta}]$ such that $\inf_{(\gamma, \delta) \in \mathcal{C}_{\hat{\beta}}} \Pi^\Phi(\gamma, \delta) = \Pi^\Phi(\gamma^*(\hat{\beta}), \hat{\beta})$. Now we get that $(\alpha Y, Z \cdot X^c + U \star \tilde{\mu} + N) \in \mathbb{H}_{\hat{\beta}}^2 \times \mathcal{H}_{\hat{\beta}}^2$, and due to the orthogonality of the martingales we conclude that $(\alpha Y, Z, U, N) \in \mathbb{H}_{\hat{\beta}}^2 \times \mathbb{H}_{\hat{\beta}}^{2,c} \times \mathbb{H}_{\hat{\beta}}^{2,d} \times \mathcal{H}_{\hat{\beta}}^{2,\perp}$. Hence, the operator

$$S : \mathbb{H}_{\hat{\beta}}^2 \times \mathbb{H}_{\hat{\beta}}^{2,c} \times \mathbb{H}_{\hat{\beta}}^{2,d} \times \mathcal{H}_{\hat{\beta}}^{2,\perp} \longrightarrow \mathbb{H}_{\hat{\beta}}^2 \times \mathbb{H}_{\hat{\beta}}^{2,c} \times \mathbb{H}_{\hat{\beta}}^{2,d} \times \mathcal{H}_{\hat{\beta}}^{2,\perp},$$

with the associated norms, that maps the processes $(\alpha y, z, u, n)$ to the processes $(\alpha Y, Z, U, N)$ defined above, is indeed well-defined.

Let $(\alpha y^i, z^i, u^i, n^i) \in \mathbb{H}_{\hat{\beta}}^2 \times \mathbb{H}_{\hat{\beta}}^{2,c} \times \mathbb{H}_{\hat{\beta}}^{2,d} \times \mathcal{H}_{\hat{\beta}}^{2,\perp}$ for $i = 1, 2$, with

$$S(\alpha y^i, z^i, u^i, n^i) = (\alpha Y^i, Z^i, U^i, N^i), \quad \text{for } i = 1, 2.$$

Denote, as usual, $\delta y, \delta z, \delta u, \delta n$ the difference of the processes and $\psi_t := f(t, y_t^1, z_t^1, u_t^1(\cdot)) - f(t, y_t^2, z_t^2, u_t^2(\cdot))$. It is immediate that $\frac{\psi}{\alpha} \in \mathbb{H}_{\hat{\beta}}^2$ and that

$$\begin{aligned} \|S(\alpha y^1, z^1, u^1, n^1) - S(\alpha y^2, z^2, u^2, n^2)\|_{\hat{\beta}}^2 &= \|\alpha \delta Y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta U\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 + \|\delta N\|_{\mathcal{H}_{\hat{\beta}}^{2,\perp}}^2 \\ &\stackrel{\delta \xi=0}{\leq} \stackrel{\text{Lem. 3.7}}{M^\Phi(\hat{\beta})} \left\| \frac{\psi}{\alpha} \right\|_{\mathbb{H}_{\hat{\beta}}^2}^2 \\ &\stackrel{(3.29)}{\leq} 2M^\Phi(\hat{\beta}) \left(\|\alpha \delta y\|_{\mathbb{H}_{\hat{\beta}}^2}^2 + \|\delta z\|_{\mathbb{H}_{\hat{\beta}}^{2,c}}^2 + \|\delta u\|_{\mathbb{H}_{\hat{\beta}}^{2,d}}^2 \right) \\ &\leq 2M^\Phi(\hat{\beta}) \|(\alpha y^1, z^1, u^1, n^1) - (\alpha y^2, z^2, u^2, n^2)\|_{\hat{\beta}}^2. \end{aligned}$$

Hence, for $M^\Phi(\hat{\beta}) \leq 1/2$, we can apply Banach's fixed point theorem to obtain the existence of a unique fixed point (\tilde{Y}, Z, U, N) . To obtain a solution in the desirable spaces, we substitute \tilde{Y} in the triplet with Y , the corresponding càdlàg version; indeed, \mathbb{G} satisfies the usual conditions and \tilde{Y} is a semimartingale.

The exact same reasoning using the $\|\cdot\|_{\mathcal{S}_\beta^2}$ -norm for Y leads to a contraction when $M_\star^\Phi(\hat{\beta}) < 1/2$. \square

Corollary 3.13 (Picard approximation). *Assume that $M^\Phi(\hat{\beta}) < 1/2$ (resp. $M_\star^\Phi(\hat{\beta}) < 1/2$) and define a sequence $(\Upsilon^{(p)})_{p \in \mathbb{N}}$ on $\mathbb{H}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$ (resp. on $\mathcal{S}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$) such that $\Upsilon^{(0)}$ is the zero element of the product space and $\Upsilon^{(p+1)}$ is the solution of*

$$\begin{aligned} Y_t^{(p+1)} = & \xi + \int_t^T f(s, Y_s^{(p)}, Z_s^{(p)}, U_s^{(p)}(\cdot)) dC_s - \int_t^T Z_s^{(p+1)} dX_s^c - \int_t^T dN_s^{(p+1)} \\ & - \int_t^T \int_{\mathbb{R}^n} U_s^{(p+1)}(x) \tilde{\mu}(ds, dx) \end{aligned}$$

Then

(i) *The sequence $(\Upsilon^{(p)})_{p \in \mathbb{N}}$ converges in $\mathbb{H}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$ (resp. in $\mathcal{S}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$) to the solution of the BSDE (3.5).*

(ii) *The following convergence holds*

$$(Z^{(p)}, U_s^{(p)}, N^{(p)}) \xrightarrow{p \rightarrow \infty} (Z, U, N), \text{ in } \mathbb{H}_\beta^2(X^c) \times \mathbb{H}_\beta^2(X^d) \times \mathbb{H}_\beta^2(X^\perp).$$

(iii) *There exists a subsequence $(\Upsilon^{(p_m)})_{m \in \mathbb{N}}$ which converges $e^{\hat{\beta}A_t} dC_t \otimes d\mathbb{P} - a.e.$*

Proof. As in the proof of Theorem 3.5, we obtain

$$\|\Upsilon^{(p+1)} - \Upsilon^{(p)}\|_{\hat{\beta}}^2 \leq (2M^\Phi(\hat{\beta}))^p \|\Upsilon^{(1)}\|_{\hat{\beta}}^2 \quad \left(\text{resp. } \|\Upsilon^{(p+1)} - \Upsilon^{(p)}\|_{\star, \hat{\beta}}^2 \leq (2M_\star^\Phi(\hat{\beta}))^p \|\Upsilon^{(1)}\|_{\star, \hat{\beta}}^2 \right), \quad (3.32)$$

and consequently, since $\sum_{p \in \mathbb{N}} \|\Upsilon^{(p+1)} - \Upsilon^{(p)}\|_{\hat{\beta}}^2 < \infty$ (resp. $\sum_{p \in \mathbb{N}} \|\Upsilon^{(p+1)} - \Upsilon^{(p)}\|_{\star, \hat{\beta}}^2 < \infty$), the sequence $(\Upsilon^{(p)})_{p \in \mathbb{N}}$ is Cauchy in $\mathbb{H}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$ (resp. in $\mathcal{S}_\beta^2 \times \mathbb{H}_\beta^{2,c} \times \mathbb{H}_\beta^{2,d} \times \mathcal{H}_\beta^{2,\perp}$). Denote by Υ the unique limit on the product space. Then, it coincides with the unique fixed point for the contraction S (see the proof of Theorem 3.5 above) due to the construction of $(\Upsilon^{(p)})_{p \in \mathbb{N}}$, which proves (i).

For (ii), the result is immediate by the Cauchy property of the sequence $(\Upsilon^{(p)})_{p \in \mathbb{N}}$, Itô's isometry, the stability of the closed linear space generated by X , which makes X^\perp to be a closed subspace, see [65, Theorem 6.16], and the orthogonality of X^c and X^d .

Finally, for (iii), by the $\|\cdot\|_{\hat{\beta}}$ -convergence, we can extract a subsequence $\{p_m\}_{m \in \mathbb{N}}$ such that

$$\|\Upsilon^{(p_{m+1})} - \Upsilon^{(p_m)}\|_{\hat{\beta}} \leq 2^{-2m}, \text{ for every } m \geq 0. \quad (3.33)$$

Define, for any $\varepsilon \geq 0$, $N^{p, \varepsilon} := \left\{ (\omega, t) \in \Omega \times [0, T], |Y_t^{(p)}(\omega) - Y_t(\omega)| > \varepsilon \right\}$. Then we have

$$\begin{aligned} e^{\hat{\beta}A_t} dC_t \otimes d\mathbb{P} \left(\limsup_{m \rightarrow \infty} N^{p_m, \varepsilon} \right) &= \lim_{m \rightarrow \infty} e^{\hat{\beta}A_t} dC_t \otimes d\mathbb{P} \left(\bigcup_{\ell=m}^{\infty} [|Y^{(p_\ell)} - Y| > \varepsilon] \right) \\ &\leq \lim_{m \rightarrow \infty} \frac{1}{\varepsilon^2} \sum_{\ell=m}^{\infty} \mathbb{E} \left[\int_0^T e^{\hat{\beta}A_t} |Y_t^{(p_\ell)} - Y_t|^2 dC_t \right] \\ &\leq \lim_{m \rightarrow \infty} \frac{1}{\varepsilon^2} \sum_{\ell=m}^{\infty} \|Y^{(p_\ell)} - Y\|_{\hat{\beta}}^2 \\ &\leq \lim_{m \rightarrow \infty} \frac{1}{\varepsilon^2} \sum_{\ell=m}^{\infty} \left(\sum_{n=1}^{\infty} 2^n \|Y^{(p_{\ell+n+1})} - Y^{(p_{\ell+n})}\|_{\mathbb{H}_\beta^2}^2 \right) \end{aligned}$$

$$\begin{aligned}
&\leq \lim_{m \rightarrow \infty} \frac{1}{\varepsilon^2} \sum_{m=\ell}^{\infty} \left(\sum_{n=1}^{\infty} 2^n \left\| \Upsilon^{(p_{\ell+n+1})} - \Upsilon^{(p_{\ell+n})} \right\|_{\hat{\beta}}^2 \right) \\
&\stackrel{(3.33)}{\leq} \lim_{m \rightarrow \infty} \frac{1}{\varepsilon^2} \sum_{\ell=m}^{\infty} \left(\sum_{n=1}^{\infty} 2^n 2^{-2(\ell+n)} \right) = 0, \text{ for any } \varepsilon > 0.
\end{aligned}$$

Hence

$$e^{\hat{\beta} A_t} dC_t \otimes d\mathbb{P} \left(\limsup_{m \rightarrow \infty} N^{p_m, 0} \right) \leq \sum_{n \in \mathbb{N}} e^{\hat{\beta} A_t} dC_t \otimes d\mathbb{P} \left(\limsup_{m \rightarrow \infty} N^{p_m, 1/n} \right) = 0.$$

Following the same arguments, we have the almost sure convergence of $Z^{p_m}, U^{p_m}, N^{p_m}$ to the corresponding processes of the $\|\cdot\|_{\beta}$ -solution of the BSDE (3.5). Moreover, using the same steps, we can obtain the analogous result for the $\|\cdot\|_{\star, \hat{\beta}}$ -norm. \square

4. APPLICATIONS

As an application of the main theorem, we show that a BSDE driven by an extended Grigelionis process, which is, roughly speaking, a superposition of a time-inhomogeneous Lévy process with a (discrete-time) random walk, admits a unique solution under appropriate conditions. The main point here is that when C is allowed to have jumps, there is a subtle interplay between the size of the jumps of C and the strength of the dependence of the generator of the BSDE, measured by the value of the Lipschitz coefficients, in the sense that their product has to remain small.

Definition 4.1. A square-integrable \mathbb{R}^m -valued martingale X is called *K-almost quasi-left-continuous* if there exists a constant $K \geq 0$ such that $|\Delta \langle X^{i,j} \rangle|_t \leq K$ for every $i, j = 1, \dots, \ell$ and for every $t \in \mathbb{R}_+, \mathbb{P} - a.s.$ In other words, the predictable quadratic covariation $\langle X \rangle$ of X has jumps uniformly bounded by K .

The next result follows directly from the definition above and Theorem 3.5.

Corollary 4.2. Let $(X, \mathbb{G}, T, \xi, f, C)$ be standard data under $\hat{\beta}$, X be *K-almost quasi-left-continuous* and the process α^2 (defined in (F4)) be bounded by $1/(18emK)$, $\mathbb{P} - a.s.$, where m is the dimension of X . Then, for $C = \text{Tr}[\langle X \rangle]$ and for $\hat{\beta}$ large enough, there exists a unique solution (Y, Z, U, N) to the BSDE (3.5).

Example 4.3. Let $(X, \mathbb{G}, T, \xi, f, C)$ be standard data under $\hat{\beta}$ such that $X = \lambda G$, for some $\lambda \in \mathbb{R}$ and some extended Grigelionis martingale G . In other words, C can be chosen to be of the form

$$C_t = \lambda^2 \left(t + \sum_{s \leq t} \mathbb{1}_{\Theta}(s) \right),$$

where $\Theta \subset (0, +\infty)$ is at most countable, see [75, Definition 2.15]. Then, since X is λ^2 -almost quasi-left-continuous, for α^2 bounded by $1/(18e\lambda^2)$ and for $\hat{\beta}$ large enough, there exists a unique solution to the BSDE (3.5).

Another interesting application of this result consists in ensuring the existence and uniqueness of the solution of the BSDE (3.5) when X is the continuous-time extension of a discrete time martingale \hat{X} . In particular, when \hat{X}^n is the discretisation of a square integrable, quasi-left continuous martingale with independent increments. Then, as the mesh of the grid tends to 0, the bound K^n of the jumps of $\langle \hat{X}^n \rangle$ tends to 0. Hence, we have that the sequence of BSDEs is, for n large enough, well-posed, given that the associated α^2 is bounded $\mathbb{P} - a.s.$

APPENDIX A. PROOF OF LEMMA 2.5

The proof of Lemma 2.5 heavily relies on Lemma A.1.(vii), which is a complement result to the already known ones about generalised inverses.

Lemma A.1. *Let $A : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \cup \{\infty\}$ be a càdlàg and increasing function with $A_0 = 0$. Denote by $L : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \cup \{\infty\}$ the càglàd generalized inverse of A , i.e.*

$$L(s) := \inf \{t \in \mathbb{R}_+, A(t) \geq s\}$$

and by $R : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \cup \{\infty\}$ the càdlàg generalized inverse of A , i.e.

$$R(s) := \inf \{t \in \mathbb{R}_+, A(t) > s\}.$$

We have

- (i) L, R are increasing.
- (ii) $L(s) = R(s-)$ and $L(s+) = R(s)$.
- (iii) $s \leq A(t)$ if and only if $L(s) \leq t$ and $s < A(t)$ if and only if $R(s) < t$.
- (iv) $A(t) < s$ if and only if $t < L(s)$ and $A(t) \leq s$ if and only if $t \leq R(s)$.
- (v) $A(R(s)) \geq A(L(s)) \geq s$, for $s \in \mathbb{R}_+$, and at most one of the inequalities can be strict.
- (vi) For $s \in A(\mathbb{R}_+)$, $A(L(s)) = s$.
- (vii) For s such that $L(s) < \infty$, we have

$$s \leq A(L(s)) \leq s + \Delta A(L(s)),$$

where $\Delta A(L(s))$ is the jump of the function A at the point $L(s)$.

Proof. Here we will prove only inequality (vii). However, we present the other properties, since we will make use of them in the proof. The interested reader can find their proofs in a slightly more general framework in [55] and the references therein.

We need to prove that $A(L(s)) - s \leq \Delta A(L(s))$ for any s such that $L(s) < \infty$. By (vi), when $s \in A(\mathbb{R}_+)$, we have since A is increasing

$$A(L(s)) - s = 0 \leq \Delta A(L(s)).$$

Now if $s \notin A(\mathbb{R}_+)$ and $s > A_\infty$, then $L(s) = \infty$, so that this case is automatically excluded. Therefore, we now assume that $s \notin A(\mathbb{R}_+)$ and $s \leq A_\infty$. Since $s \notin A(\mathbb{R}_+)$, there exists some $t \in \mathbb{R}_+$ such that $s \in [A(t-), A(t))$. Then, we immediately have $L(s) = t$. Hence

$$s + \Delta A(L(s)) = s + \Delta A(t) \geq A(t) = A(L(s)),$$

since $s \geq A(t-)$. □

Proof of Lemma 2.5. Using a change of variables, Lemma A.1.(vii) and that g is non-decreasing and submultiplicative, we have

$$\begin{aligned} \int_0^t g(A_s) dA_s &= \int_{A_0}^{A_t} g(A_{L_s}) ds \leq \int_{A_0}^{A_t} g(s + \Delta A_{L_s}) ds \\ &\leq \int_{A_0}^{A_t} g\left(s + \max_{\{s, L_s < \infty\}} \Delta A_{L_s}\right) ds \leq cg\left(\max_{\{s, L_s < \infty\}} \Delta A_{L_s}\right) \int_{A_0}^{A_t} g(s) ds. \end{aligned}$$

□

APPENDIX B. PROOF OF LEMMA 3.4

Proof. Let $(\gamma, \delta) \in \mathcal{C}_\beta$. We shall begin by obtaining the critical points of the map Π^Φ . We have

$$\begin{aligned} \frac{\partial}{\partial \gamma} \Pi^\Phi(\gamma, \delta) &= (2 + 9\delta) e^{(\delta-\gamma)\Phi} \frac{\Phi \gamma^2 + (2 - \delta\Phi)\gamma - \delta}{\gamma^2(\delta - \gamma)^2}, \\ \frac{\partial}{\partial \delta} \Pi^\Phi(\gamma, \delta) &= -\frac{9}{\delta^2} + e^{(\delta-\gamma)\Phi} \left\{ \frac{[9 + (2 + 9\delta)\Phi](\delta - \gamma)}{\gamma(\delta - \gamma)^2} - \frac{2 + 9\delta}{\gamma(\delta - \gamma)^2} \right\}. \end{aligned}$$

The only possible critical points for Π^Φ are therefore such that $\delta = -2/9$ or $\gamma = \frac{\delta\Phi - 2 + \sqrt{4 + \delta^2\Phi^2}}{2\Phi}$. However, the values $\delta = -2/9$ and $\gamma = \frac{\delta\Phi - 2 - \sqrt{4 + \delta^2\Phi^2}}{2\Phi}$ are ruled out as negative. For $0 < \delta \leq \beta$ we have

$$\left(\frac{\delta\Phi - 2 + \sqrt{4 + \delta^2\Phi^2}}{2\Phi}, \delta \right) \in \mathcal{C}_\beta.$$

Let us define $\bar{\gamma}^\Phi(\delta) := \frac{\delta\Phi - 2 + \sqrt{4 + \delta^2\Phi^2}}{2\Phi}$, for $0 < \delta \leq \beta$. It is easy to verify that $\bar{\gamma}^\Phi(\delta) \in (0, \delta)$. Then, some tedious calculations yield that

$$\frac{\partial \Pi^\Phi}{\partial \delta}(\bar{\gamma}^\Phi(\delta), \delta) = -\frac{9}{\delta^2} - \frac{\exp[(\delta - \bar{\gamma}^\Phi(\delta))\Phi]}{\bar{\gamma}^\Phi(\delta)(\delta - \bar{\gamma}^\Phi(\delta))^2} \cdot \frac{2\bar{\gamma}^\Phi(\delta)\Phi + 9\bar{\gamma}^\Phi(\delta) + 2}{(\bar{\gamma}^\Phi(\delta)\Phi + 1)^2} < 0$$

therefore Π^Φ does not admit any critical point on \mathcal{C}_β , for which $0 < \gamma < \delta < \beta$. Hence, the infimum on this set is necessarily attained on its boundary. The cases where at least one among δ and γ goes to 0, or where their difference goes to 0, lead to the value ∞ . The only remaining case is therefore $0 < \gamma < \delta = \beta$, where β is fixed. Then we get

$$\frac{d}{d\gamma} \Pi^\Phi(\gamma, \beta) = (2 + 9\beta) e^{(\beta - \gamma)\Phi} \frac{\Phi\gamma^2 + (2 - \beta\Phi)\gamma - \beta}{\gamma^2(\beta - \gamma)^2},$$

and $\Pi^\Phi(\gamma, \beta)$ viewed as a function of γ attains its minimum at its critical point given by $\bar{\gamma}^\Phi(\beta)$, since $\frac{d\Pi^\Phi}{d\gamma}(\gamma, \beta) < 0$ on $(0, \bar{\gamma}^\Phi(\beta))$ and $\frac{d\Pi^\Phi}{d\gamma}(\gamma, \beta) > 0$ on $(\bar{\gamma}^\Phi(\beta), \beta)$.

Now, we proceed to the second case, and start by determining the critical points of Π_\star^Φ . It holds

$$\begin{aligned} \frac{\partial}{\partial \gamma} \Pi_\star^\Phi(\gamma, \delta) &= -\frac{8}{\gamma^2} + 9\delta e^{(\delta - \gamma)\Phi} \frac{\Phi\gamma^2 - (\delta\Phi - 2)\gamma - \delta}{\gamma^2(\delta - \gamma)^2}, \\ \frac{\partial}{\partial \delta} \Pi_\star^\Phi(\gamma, \delta) &= -\frac{9}{\delta^2} + 9e^{(\delta - \gamma)\Phi} \frac{(1 + \delta\Phi)(\delta - \gamma) - \delta}{\gamma(\delta - \gamma)^2}. \end{aligned}$$

Following analogous computations as above we can prove that, for $(\gamma, \delta) \in \mathcal{C}_\beta$, the equation

$$\frac{\partial}{\partial \gamma} \Pi_\star^\Phi(\gamma, \delta) = 0 \Leftrightarrow P_\delta(\gamma) := 8(\delta - \gamma)^2 - 9\delta e^{(\delta - \gamma)\Phi} (\Phi\gamma^2 - (\delta\Phi - 2)\gamma - \delta) = 0$$

has a unique root, say $\bar{\gamma}_\star^\Phi(\delta)$, which moreover satisfies $\bar{\gamma}_\star^\Phi(\delta) \in (\bar{\gamma}^\Phi(\delta), \delta)$. This can be proved because the function $P_\delta : (0, \delta) \rightarrow \mathbb{R}$ is decreasing, for each fixed $\delta \in (0, \beta)$, with $P_\delta(\bar{\gamma}^\Phi(\delta)) > 0$ and $P_\delta(\delta) < 0$. Now observe that for $\gamma > \frac{\delta^2\Phi}{1 + \delta\Phi}$ it holds $\frac{\partial}{\partial \delta} \Pi_\star^\Phi(\gamma, \delta) < 0$ and that $P_\delta(\frac{\delta^2\Phi}{1 + \delta\Phi}) > 0$. Using the monotonicity of P_δ we have that $\bar{\gamma}_\star^\Phi(\delta) > \frac{\delta^2\Phi}{1 + \delta\Phi}$ and therefore also $\frac{\partial}{\partial \delta} \Pi_\star^\Phi(\bar{\gamma}_\star^\Phi(\delta), \delta) < 0$. Arguing as above we can conclude that the infimum is attained for $\delta = \beta$ at the point $(\bar{\gamma}_\star^\Phi(\beta), \beta)$.

Finally, the limiting statements follow by straightforward but tedious computations. \square

APPENDIX C. AUXILIARY RESULTS ON OPTIONAL MEASURES

Let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be a filtered probability space and $Y = \{Y_t\}_{t \in [0, \infty]}$ be a uniformly integrable measurable process. Then, thanks to the uniform integrability, we have by a clear adaptation of [65, Theorem 5.1] that there exists a unique optional process, denoted by oY , such that for every \mathbb{G} -stopping time τ , we have $\mathbb{E}_\tau[Y_\tau] = {}^oY_\tau$, \mathbb{P} -a.s. Observe that τ is allowed to take infinite values, since Y_∞ is well-defined and integrable. For any increasing, càdlàg and \mathbb{G} -adapted process A , the measure $\mu_A : (\Omega \times [0, \infty], \mathcal{G} \otimes \mathcal{B}([0, \infty])) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ defined as

$$\mu_A(H) = \mathbb{E} \left[\int_0^\infty \mathbf{1}_H dA_t \right], \text{ for } H \in \mathcal{G} \otimes \mathcal{B}([0, \infty]),$$

is optional, see [65, Definition 5.10, Definition 5.12 and Theorem 5.13]. For convenience, we state the following well-known result as a lemma.

Lemma C.1. *Let A be an increasing, càdlàg and adapted process and Y be a uniformly integrable and measurable process, then it holds $\mu_A(Y) = \mu_A({}^oY)$.*

Proof. Let L be the càglàd generalised inverse of A (see Lemma 2.5 for the definition). We have

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty Y_t dA_t\right] &= \mathbb{E}\left[\int_0^\infty Y_{L_s} \mathbb{1}_{[L_s < \infty]} ds\right] = \int_0^\infty \mathbb{E}[Y_{L_s} \mathbb{1}_{[L_s < \infty]}] ds \\ &= \int_0^\infty \mathbb{E}[{}^oY_{L_s} \mathbb{1}_{[L_s < \infty]}] ds = \mathbb{E}\left[\int_0^\infty {}^oY_t dA_t\right],\end{aligned}$$

where for the change of variables we used [65, Lemma 1.38], and for the third equality the definition of the optional projection, see [65, Theorem 5.1]. \square

REFERENCES

- [1] V.I. Arkin and M.T. Saksonov. Necessary conditions of optimality in the problems of control of stochastic differential equations. *Doklady Akademii Nauk SSSR*, 244(1):11–15, 1979.
- [2] K. Bahlali. Backward stochastic differential equations with locally Lipschitz coefficient. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 333(5):481–486, 2001.
- [3] K. Bahlali. Existence and uniqueness of solutions for BSDEs with locally Lipschitz coefficient. *Electronic Communications in Probability*, 7:169–179, 2002.
- [4] K. Bahlali, M. Eddahbi, and E.H. Essaky. BSDE associated with Lévy processes and application to PDIE. *International Journal of Stochastic Analysis*, 16(1):1–17, 2003.
- [5] K. Bahlali, M. Eddahbi, and Y. Ouknine. Solvability of some quadratic BSDEs without exponential moments. *Comptes Rendus Mathématique*, 351(5):229–233, 2013.
- [6] K. Bahlali and B. El Asri. Stochastic optimal control and BSDEs with logarithmic growth. *Bulletin des Sciences Mathématiques*, 136(6):617–637, 2012.
- [7] K. Bahlali, E.H. Essaky, and M. Hassani. Multidimensional BSDEs with super-linear growth coefficient: application to degenerate systems of semilinear PDEs. *Comptes Rendus Mathématique*, 348(11):677–682, 2010.
- [8] K. Bahlali, E.H. Essaky, and M. Hassani. Existence and uniqueness of multidimensional BSDEs and of systems of degenerate PDEs with superlinear growth generator. *SIAM Journal on Mathematical Analysis*, 47(6):4251–4288, 2015.
- [9] K. Bahlali, E.H. Essaky, M. Hassani, and É. Pardoux. Existence, uniqueness and stability of backward stochastic differential equations with locally monotone coefficient. *Comptes Rendus Mathématique*, 335(9):757–762, 2002.
- [10] E. Bandini. Existence and uniqueness for backward stochastic differential equations driven by a random measure, possibly non quasi-left continuous. *Electronic Communications in Probability*, 20, 2015.
- [11] G. Barles, R. Buckdahn, and É. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 60(1-2):57–83, 1997.
- [12] P. Barrieu and N. El Karoui. Monotone stability of quadratic semimartingales with applications to unbounded general quadratic BSDEs. *The Annals of Probability*, 41(3B):1831–1863, 2013.
- [13] D. Becherer. Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging. *The Annals of Applied Probability*, 16(4):2027–2054, 2006.
- [14] C. Bender and M. Kohlmann. BSDEs with stochastic Lipschitz condition. Technical report, Center of Finance and Econometrics, University of Konstanz, 2000.
- [15] A. Bensoussan. Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions. *Stochastics: An International Journal of Probability and Stochastic Processes*, 9(3):169–222, 1983.
- [16] J.-M. Bismut. *Analyse Convexe et Probabilités*. PhD thesis, Faculté des Sciences de Paris, 1973.
- [17] J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.
- [18] J.-M. Bismut. Contrôle des systèmes linéaires quadratiques: applications de l'intégrale stochastique. *Séminaire de Probabilités de Strasbourg*, XII:180–264, 1978.
- [19] O. Bobrovnytska and M. Schweizer. Mean-variance hedging and stochastic control: beyond the Brownian setting. *IEEE Transactions on Automatic Control*, 49(3):396–408, 2004.
- [20] G. Bordigoni, A. Matoussi, and M. Schweizer. A stochastic control approach to a robust utility maximization problem. In *Stochastic Analysis and its Applications*, volume 2 of *Abel Symposia*, pages 125–151. Springer, 2007.
- [21] B. Bouchard, D. Possamaï, X. Tan, and C. Zhou. A unified approach to *a priori* estimates for supersolutions of BSDEs in general filtrations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques (B)*, to appear, 2015.
- [22] P. Briand and R. Carmona. BSDEs with polynomial growth generators. *International Journal of Stochastic Analysis*, 13(3):207–238, 2000.
- [23] P. Briand and F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochastic Processes and their Applications*, 118(5):818–838, 2008.
- [24] P. Briand and F. Confortola. Quadratic BSDEs with random terminal time and elliptic PDEs in infinite dimension. *Electronic Journal of Probability*, 13(54):1529–1561, 2008.
- [25] P. Briand, B. Delyon, Y. Hu, É. Pardoux, and L. Stoica. L^p solutions of backward stochastic differential equations. *Stochastic Processes and their Applications*, 108(1):109–129, 2003.

- [26] P. Briand, B. Delyon, and J. Mémin. On the robustness of backward stochastic differential equations. *Stochastic Processes and their Applications*, 97(2):229–253, 2002.
- [27] P. Briand and R. Élie. A simple constructive approach to quadratic BSDEs with or without delay. *Stochastic Processes and their Applications*, 123(8):2921–2939, 2013.
- [28] P. Briand and Y. Hu. Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs. *Journal of Functional Analysis*, 155(2):455–494, 1998.
- [29] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probability Theory and Related Fields*, 136(4):604–618, 2006.
- [30] P. Briand and Y. Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probability Theory and Related Fields*, 141(3-4):543–567, 2008.
- [31] P. Briand, J.-P. Lepeltier, and J. San Martín. One-dimensional backward stochastic differential equations whose coefficient is monotonic in y and non-Lipschitz in z . *Bernoulli*, 13(1):80–91, 2007.
- [32] R. Buckdahn. Backward stochastic differential equations driven by a martingale. *preprint*, 1993.
- [33] R. Buckdahn and É. Pardoux. BSDE's with jumps and associated integro-partial differential equations. *preprint*, 1994.
- [34] R. Carbone, B. Ferrario, and M. Santacroce. Backward stochastic differential equations driven by càdlàg martingales. *Theory of Probability & its Applications*, 52(2):304–314, 2008.
- [35] P. Cheridito and K. Nam. BSDEs with terminal conditions that have bounded Malliavin derivative. *Journal of Functional Analysis*, 266(3):1257–1285, 2014.
- [36] P. Cheridito and K. Nam. Multidimensional quadratic and subquadratic BSDEs with special structure. *Stochastics: An International Journal of Probability and Stochastic Processes*, 87(5):871–884, 2015.
- [37] P. Cheridito and M. Stadje. Existence, minimality and approximation of solutions to BSDEs with convex drivers. *Stochastic Processes and their Applications*, 122(4):1540–1565, 2012.
- [38] R.J. Chitashvili. Martingale ideology in the theory of controlled stochastic processes. In *Probability Theory and Mathematical Statistics*, pages 73–92. Springer, 1983.
- [39] R.J. Chitashvili and M.G. Mania. Optimal locally absolutely continuous change of measure. Finite set of decisions. Part I. *Stochastics: An International Journal of Probability and Stochastic Processes*, 21(2):131–185, 1987.
- [40] R.J. Chitashvili and M.G. Mania. Optimal locally absolutely continuous change of measure. Finite set of decisions. Part II: optimization problems. *Stochastics: An International Journal of Probability and Stochastic Processes*, 21(3):187–229, 1987.
- [41] S.N. Cohen and R.J. Elliott. Existence, uniqueness and comparisons for BSDEs in general spaces. *The Annals of Probability*, 40(5):2264–2297, 2012.
- [42] S.N. Cohen, R.J. Elliott, and C.E.M. Pearce. A general comparison theorem for backward stochastic differential equations. *Advances in Applied Probability*, 42(3):878–898, 2010.
- [43] F. Confortola, M. Fuhrman, and J. Jacod. Backward stochastic differential equation driven by a marked point process: an elementary approach with an application to optimal control. *The Annals of Applied Probability*, 26:1743–1773, 2016.
- [44] R.W.R. Darling and É. Pardoux. Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *The Annals of Probability*, 25(3):1135–1159, 1997.
- [45] M.H.A. Davis and P. Varaiya. Dynamic programming conditions for partially observable stochastic systems. *SIAM Journal on Control and Optimization*, 11(2):226–261, 1973.
- [46] M.H.A. Davis and P. Varaiya. The multiplicity of an increasing family of σ -fields. *The Annals of Probability*, 2(5):958–963, 1974.
- [47] F. Delbaen, Y. Hu, and X. Bao. Backward SDEs with superquadratic growth. *Probability Theory and Related Fields*, 150(1-2):145–192, 2011.
- [48] F. Delbaen, Y. Hu, and A. Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques (B)*, 47(2):559–574, 2011.
- [49] F. Delbaen, Y. Hu, and A. Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions: the critical case. *Discrete and Continuous Dynamical Systems-Series A*, 35(11):5273–5283, 2015.
- [50] S. Drapeau, G. Heyne, and M. Kupper. Minimal supersolutions of convex BSDEs. *The Annals of Probability*, 41:3973–4001, 2013.
- [51] N. El Karoui and S.-J. Huang. A general result of existence and uniqueness of backward stochastic differential equations. In N. El Karoui and L. Mazliak, editors, *Backward stochastic differential equations*, volume 364 of *Chapman & Hall/CRC Research Notes in Mathematics Series*, pages 27–36. Longman, 1997.
- [52] N. El Karoui, A. Matoussi, and A. Ngoupeyou. Quadratic exponential semimartingales and application to BSDEs with jumps. *arXiv preprint arXiv:1603.06191*, 2016.
- [53] N. El Karoui, S. Peng, and M.-C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.
- [54] R. Élie and D. Possamaï. Contracting theory with competitive interacting agents. *arXiv preprint arXiv:1605.08099*, 2016.

- [55] P. Embrechts and M. Hofert. A note on generalized inverses. *Mathematical Methods of Operations Research*, 77(3):423–432, 2013.
- [56] G.-E. Espinosa and N. Touzi. Optimal investment under relative performance concerns. *Mathematical Finance*, 25(2):221–257, 2015.
- [57] S.-J. Fan. Bounded solutions, L^p ($p > 1$) solutions and L^1 solutions for one dimensional BSDEs under general assumptions. *Stochastic Processes and their Applications*, 126:1511–1552, 2016.
- [58] C. Frei. Splitting multidimensional BSDEs and finding local equilibria. *Stochastic Processes and their Applications*, 124(8):2654–2671, 2014.
- [59] C. Frei and G. dos Reis. A financial market with interacting investors: does an equilibrium exist? *Mathematics and Financial Economics*, 4(3):161–182, 2011.
- [60] M. Fujii and A. Takahashi. Quadratic-exponential growth BSDEs with jumps and their Malliavin’s differentiability. *arXiv preprint arXiv:1512.05924*, 2015.
- [61] S. Hamadène. Équations différentielles stochastiques rétrogrades: le cas localement lipschitzien. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques (B)*, 32(5):645–659, 1996.
- [62] S. Hamadène, J.-P. Lepeltier, and S. Peng. BSDEs with continuous coefficients and stochastic differential games. In N. El Karoui and L. Mazliak, editors, *Backward stochastic differential equations*, volume 364 of *Chapman & Hall/CRC Research Notes in Mathematics Series*, pages 115–128. Longman, 1997.
- [63] J. Harter and A. Richou. A stability approach for solving multidimensional quadratic BSDEs. *arXiv preprint arXiv:1606.08627*, 2016.
- [64] M. Hassani and Y. Ouknine. On a general result for backward stochastic differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 73(3-4):219–240, 2002.
- [65] S. He, J. Wang, and J.A. Yan. *Semimartingale Theory and Stochastic Calculus*. Science Press, 1992.
- [66] Y. Hu and M. Schweizer. Some new BSDE results for an infinite-horizon stochastic control problem. In *Advanced Mathematical Methods for Finance*, pages 367–395. Springer, 2011.
- [67] Y. Hu and S. Tang. Multi-dimensional backward stochastic differential equations of diagonally quadratic generators. *Stochastic Processes and their Applications*, 126(4):1066–1086, 2016.
- [68] Y. Hu and G. Tessitore. BSDE on an infinite horizon and elliptic PDEs in infinite dimension. *Nonlinear Differential Equations and Applications NoDEA*, 14(5-6):825–846, 2007.
- [69] P. Imkeller, A. Réveillac, and A. Richter. Differentiability of quadratic BSDEs generated by continuous martingales. *The Annals of Applied Probability*, 22(1):285–336, 2012.
- [70] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288. Springer Science & Business Media, 2013.
- [71] A. Jamneshan, M. Kupper, and P. Luo. Multidimensional quadratic BSDEs with separated generators. *arXiv preprint arXiv:1501.00461*, 2014.
- [72] A. Jamneshan, M. Kupper, and P. Luo. Solvability of multidimensional quadratic BSDEs. *arXiv preprint arXiv:1612.02698*, 2016.
- [73] M. Jeanblanc, M.G. Mania, M. Santacrose, and M. Schweizer. Mean-variance hedging via stochastic control and BSDEs for general semimartingales. *The Annals of Applied Probability*, 22(6):2388–2428, 2012.
- [74] Yu.M. Kabanov. On the Pontryagin maximum principle for the linear stochastic differential equations. *Probabilistic Models and Control of Economical Processes*, 1978.
- [75] J. Kallsen. *Semimartingale Modelling in Finance*. PhD thesis, University of Freiburg, 1998.
- [76] C. Kardaras, H. Xing, and G. Žitković. Incomplete stochastic equilibria with exponential utilities close to Pareto optimality. *arXiv preprint arXiv:1505.07224*, 2015.
- [77] N. Kazi-Tani, D. Possamaï, and C. Zhou. Quadratic BSDEs with jumps: a fixed-point approach. *Electronic Journal of Probability*, 20(66):1–28, 2015.
- [78] N. Kazi-Tani, D. Possamaï, and C. Zhou. Quadratic BSDEs with jumps: related nonlinear expectations. *Stochastics and Dynamics*, 1650012, 2015.
- [79] M. Kobylanski. Résultats d’existence et d’unicité pour des équations différentielles stochastiques rétrogrades avec des générateurs à croissance quadratique. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, 324(1):81–86, 1997.
- [80] M. Kobylanski. *Quelques applications de méthodes d’analyse non-linéaire à la théorie des processus stochastiques*. PhD thesis, Université de Tours, 1998.
- [81] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*, 28:558–602, 2000.
- [82] D.O. Kramkov and S. Pulido. Stability and analytic expansions of local solutions of systems of quadratic BSDEs with applications to a price impact model. *arXiv preprint arXiv:1410.6144*, 2014.
- [83] D.O. Kramkov and S. Pulido. A system of quadratic BSDEs arising in a price impact model. *The Annals of Applied Probability*, 26(2):794–817, 2016.
- [84] T. Kruse and A. Popier. BSDEs with monotone generator driven by Brownian and Poisson noises in a general filtration. *Stochastics An International Journal of Probability and Stochastic Processes*, pages 1–49, 2015.

- [85] M. Kupper, P. Luo, and L. Tangpi. Multidimensional Markov FBSDEs with superquadratic growth. *arXiv preprint arXiv:1505.01796*, 2015.
- [86] J.-P. Lepeltier and J. San Martín. Backward stochastic differential equations with continuous coefficient. *Statistics & Probability Letters*, 32(4):425–430, 1997.
- [87] J.-P. Lepeltier and J. San Martín. Existence for BSDE with superlinear-quadratic coefficient. *Stochastics: An International Journal of Probability and Stochastic Processes*, 63(3-4):227–240, 1998.
- [88] J.-P. Lepeltier and J. San Martín. On the existence or non-existence of solutions for certain backward stochastic differential equations. *Bernoulli*, 8(1):123–137, 2002.
- [89] M.G. Mania, M. Santacrose, and R. Tevzadze. A semimartingale backward equation related to the p-optimal martingale measure and the lower price of a contingent claim. *Stochastic processes and related topics*, pages 189–212, 2002.
- [90] M.G. Mania, M. Santacrose, and R. Tevzadze. A semimartingale BSDE related to the minimal entropy martingale measure. *Finance and Stochastics*, 7(3):385–402, 2003.
- [91] M.G. Mania and M. Schweizer. Dynamic exponential utility indifference valuation. *The Annals of Applied Probability*, 15(3):2113–2143, 2005.
- [92] M.G. Mania and R. Tevzadze. A semimartingale Bellman equation and the variance-optimal martingale measure. *Georgian Mathematical Journal*, 7(4):765–792, 2000.
- [93] M.G. Mania and R. Tevzadze. A semimartingale backward equation and the variance-optimal martingale measure under general information flow. *SIAM Journal on Control and Optimization*, 42(5):1703–1726, 2003.
- [94] M.G. Mania and R. Tevzadze. A unified characterization of q -optimal and minimal entropy martingale measures by semimartingale backward equations. *Georgian Mathematical Journal*, 10(2):289–310, 2003.
- [95] X. Mao. Adapted solutions of backward stochastic differential equations with non-Lipschitz coefficients. *Stochastic Processes and their Applications*, 58(2):281–292, 1995.
- [96] F. Masiero and A. Richou. A note on the existence of solutions to Markovian superquadratic BSDEs with an unbounded terminal condition. *Electronic Journal of Probability*, 18(50):1–15, 2013.
- [97] M. Mocha and N. Westray. Quadratic semimartingale BSDEs under an exponential moments condition. *Séminaire de probabilités de Strasbourg*, XLIV:105–139, 2012.
- [98] M.-A. Morlais. Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem. *Finance & Stochastics*, 13(1):121–150, 2009.
- [99] D. Nualart and W. Schoutens. Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance. *Bernoulli*, 7(5):761–776, 2001.
- [100] É. Pardoux. BSDEs, weak convergence and homogenization of semilinear PDEs. In *Nonlinear Analysis, Differential Equations and Control*, pages 503–549. Springer, 1999.
- [101] É. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *System and Control Letters*, 14(1):55–61, 1990.
- [102] S. Peng. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastic Reports*, 37(1-2):61–74, 1991.
- [103] A. Réveillac. On the orthogonal component of BSDEs in a Markovian setting. *Statistics & Probability Letters*, 82(1):151–157, 2012.
- [104] A. Richou. Markovian quadratic and superquadratic BSDEs with an unbounded terminal condition. *Stochastic Processes and their Applications*, 122(9):3173–3208, 2012.
- [105] M. Royer. BSDEs with a random terminal time driven by a monotone generator and their links with PDEs. *Stochastics and Stochastic Reports*, 76(4):281–307, 2004.
- [106] M. Royer. Backward stochastic differential equations with jumps and related non-linear expectations. *Stochastic Processes and their Applications*, 116(10):1358–1376, 2006.
- [107] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.
- [108] R. Situ. On solutions of backward stochastic differential equations with jumps and applications. *Stochastic Processes and their Applications*, 66(2):209–236, 1997.
- [109] S. Tang and X. Li. Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim.*, 32(5):1447–1475, 1994.
- [110] R. Tevzadze. Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Processes and their Applications*, 118(3):503–515, 2008.
- [111] J. Wang, Q. Ran, and Q. Chen. L^p solutions of BSDEs with stochastic Lipschitz condition. *International Journal of Stochastic Analysis*, 2007(78196), 2007.
- [112] H. Xing and G. Žitković. A class of globally solvable Markovian quadratic BSDE systems and applications. *arXiv preprint arXiv:1603.00217*, 2016.

UNIVERSITÉ PARIS–DAUPHINE, PSL RESEARCH UNIVERSITY, CNRS, CEREMADE, 75016 PARIS, FRANCE.

E-mail address: possamai@ceremade.dauphine.fr

INSTITUTE OF MATHEMATICS, TU BERLIN, STRASSE DES 17. JUNI 136, 10623 BERLIN, GERMANY

E-mail address: alsapl@math.tu-berlin.de